S. Peszat, J. Zabczyk

Wstęp do teorii sterowania stochastycznego i filtracji

20 kwietnia 2017

Spis treści

1

Wstęp

1.1 Przykłady

Przyklad 1.1 (Problem inwestora (P. Samuelson portfolio problem, 1969, [17]) Niech X_n oznacza kapitał inwestora w czasie $n = 0, 1, 2, \ldots$. W każdej chwili inwestor dzieli swój majątek. Konsumuje $c_n \in [0, X_n]$, a reszte inwestuje w dwa typy walorów. W banku, na stałą stopę zysku $1+r$, deponuje

$$b_n(X_n - c_n),$$

gdzie $b_n \in [0, 1]$ jest wybraną przez inwestora liczbą, a $r \geq 0$ jest ustalone. Za pozostałe

$$(1 - b_n)(X_n - c_n)$$

inwestor kupuje akcje, których stopa zysku jest losowa i wynosi $1 + \xi_{n+1}$. Zakładamy, że $\{\xi_n\}$ jest ciągiem niezależnych zmiennych losowych o tym samym rozkładzie. W skrócie $\{\xi_n\}$ jest *iid* (ciągiem "independent identically distributed" zmiennych losowych). Oczywiście zakładamy też, że $\mathbb{P}(\xi_n \geq -1) = 1$.

Załóżmy, że inwestor działa na przedziale czasowym $0, 1, \ldots, N$. Liczbę $N < +\infty$ nazywamy *horyzontem czasowym.* Inwestor szuka strategii konsumpcji $(c_0, \ldots, c_{N-1})$, oraz strategii inwestowania $(b_0, \ldots, b_{N-1})$, które przy danym kapitale początkowym maksymalizują jego oczekiwaną satysfakcję. Satysfakcję Samuelson proponuje mierzyć wyrażeniem

$$\sum_{n=0}^{N-1} \gamma^n c_n^\alpha + \gamma^N \omega X_N^\alpha,$$

gdzie $\gamma, \alpha \in [0, 1]$ i $\omega > 0$ są zadanymi parametrami. Tak więc oczekiwana (średnia) satysfakcja wynosi

$$\mathbb{E}\left(\sum_{n=0}^{N-1} \gamma^n c_n^\alpha + \gamma^N \omega X_N^\alpha\right). \tag{1.1}$$

W 1969 roku R.K. Merton sformułował i rozwiązał odpowiednik problemu Samuelsona w czasie ciągłym.

Przyklad 1.2 (Problem rozregulowania) Wyobraźmy sobie maszyne wytwarzającą produkt, którego jakość waha się od 1 do M. Gdy maszyna jest w dobrym stanie, to prawdopodobieństwo, że produkt jest jakości

$$k \in \{1, 2, \ldots, M-1, M\}$$

wynosi $q^0(k)$. Gdy kondycja maszyny jest zła to prawdopodobieństwo to wynosi $q^1(k)$. Załóżmy, że maszyna ulega rozregulowaniu po losowym czasie σ. Na podstawie obserwacji jakości produktów mamy podjąć decyzję czy maszyna jest sprawna, czy też należy ją zatrzymać, sprawdzić i ewentualnie naprawić. Czas zatrzymania maszyny oznaczać będziemy przez τ. Celem jest zminimalizowanie funkcjonału kosztu

$$\mathcal{J}(\tau) = \mathbb{P}(\tau < \sigma) + c\,\mathbb{E}\left((\tau - \sigma) \vee 0\right).$$

Pierwszy składnik to koszt fałszywego alarmu. Drugi to kara za opóźnienie i produkcje towaru niskiej jakości.

Przyklad 1.3 (Problem sekretarki) Załóżmy, że na posadę sekretarki zgłosiło się N kandydatek. Widząc dowolne dwie kandydatki osoba rekrutująca może bezbłędnie rozstrzygnąć, która jest lepsza. Problem leży w tym, że kandydatki zjawiają się na rozmowę kwalifikacyjną jedna po drugiej i osoba rekrutująca za każdym razem musi zadecydować czy daną kandydatkę zaangażować czy nie. Celem jest maksymalizacja prawdopodobieństwa wybrania najlepszej kandydatki.

Przyklad 1.4 (Problem optymalnego rozmieszczenia) Następujący problem został rozwiązany w [19], a sformułowany przez Guzickiego. Ma on związek ze wstępnym sortowaniem baz danych. Mianowicie załóżmy, że η_n, $n = 1, \ldots, N$, jest skończonym ciągiem niezależnych zmiennych losowych o rozkładzie jednostajnym na odcinku $[0, 1]$. Mamy też N pudełek ponumerowanych od 1 do N. Realizację ciągu η_n, $n = 1, \ldots, N$, sukcesywnie rozmieszczamy w pudełkach. Niech u_{η_j} oznacza numer pudełka, który został przyporządkowany η_j. Zasady przyporządkowania są następujące:

- do jednego pudełka możemy włożyć tylko jedną η_n,
- jeżeli $\eta_j < \eta_k$ to $u_{\eta_j} < u_{\eta_k}$,
- raz rozmieszczonych wartości nie możemy przekładać i czynność przyporządkowania musimy wykonywać na bieżąco.

W ten sposób nie zawsze uda nam się rozmieścić wszystkie wartości. Celem jest znalezienie strategii, która maksymalizuje prawdopodobieństwo prawidłowego rozmieszczenia całego ciągu. Oczywistą strategią jest podzielenie odcinka na równe N-przedziałów. Ponumerowanie rosnąco tych przedziałów, a następnie umieszczanie η_1 w pudełku o numerze odcinka do którego należy η_1. Procedure tą można stosować do rozmieszczenia η_2 i tak dalej. Czy jest ona najlepsza? Okazuje się, że nie, patrz [19]. Najlepsza strategia też polega na podziale $[0,1]$ na N odcinków, ale o niejednakowych długościach. Środkowe odcinki powinny być szersze od skrajnych.

Przyklad 1.5 (Koszty samochodowe, R. Howard 1960, [11]) Podzielmy odcinek czasowy dziesięciu lat na 40 równych trzy miesięcznych okresów. Na początku każdego z okresów właściciel decyduję czy sprzedać swój samochód i kupić inny, czy odłożyć decyzję na conajmniej trzy miesiące. Niech X_n będzie wiekiem samochodu. Zakładamy, że $X_n \in \{0,1,2,\ldots,40\}$. Samochód w wieku dziesięciu lat ($X_n = 40$) musi zostać zniszczony i zastąpiony przez inny.

Niech $u \in \{0,1,\ldots,39\}$ oznacza decyzję sprzedania samochodu i kupienia nowego w wieku u oraz niech δ oznacza decyzję odłożenia kupna na co najmniej trzy miesiące. Niech p_j będzie prawdopodobieństwem, że samochód w wieku j ulegnie całkowitemu zniszczeniu w okresie trzech miesięcy. Dla $u \in U = \{0,1,\ldots,39\} \cup \{\delta\}$ mamy następujące prawdopodobieństwa przejścia:

$$p^u_{i,j} = 1, \text{ gdy } u = j, \quad i,j = 0,1,\ldots,39,$$

$$p^\delta_{i,40} = p_i, \qquad p^\delta_{i,i+1} = 1 - p_i,$$
$$p^u_{40,i} = 1 \text{ gdy } u = i, \quad i = 0,1,\ldots,39.$$

Zakładamy, że dane są koszty $q(i,u)$, $i \in E := \{0,1,\ldots,40\}$, $u \in U$ związane z podjętymi decyzjami. Mianowicie $q(i,\delta)$ jest średnim kosztem utrzymania przez 3 miesiące samochodu w wieku i, a $q(i,j)$ jest kosztem sprzedaży samochodu w wieku i i kupienie nowego w wieku j wraz z jego utrzymaniem przez 3 miesiące. Celem właściciela jest minimalizowanie wartości oczekiwanej funkcjonału kosztu, która wynosi

$$J((u_n)) = \limsup_{N\to+\infty} \frac{1}{N}\,\mathbb{E}\sum_{n=0}^{N-1} q(X_n,u_n).$$

Funkcjonał $J((u_n))$ interpretujemy jako średni koszt utrzymania samochodu na jednostkę czasu.

1.2 Warunkowa wartość oczekiwana

W tym rozdziale przypomnimy definicję i podstawowe własności warunkowej wartości oczekiwanej. Niech $(\Omega,\mathcal{F},\mathbb{P})$ będzie przestrzenią probabilistyczną, a $\mathcal{G}$ pod-σ-ciałem $\mathcal{F}$ oraz niech $X\colon \Omega \to \mathbb{R}$ będzie $\mathcal{F}$-mierzalną zmienną losową. Będziemy zakładali, że istnieje wartość oczekiwana $\mathbb{E}\,X$, to znaczy, że $\mathbb{E}\,X^+ < +\infty$ lub $\mathbb{E}\,X^- < +\infty$.

Definicja 1.1 *Warunkową wartością oczekiwaną X względem $\mathcal{G}$* jest dowolna $\mathcal{G}$-mierzalna zmienna losowa η, dla której określona jest wartość oczekiwana $\mathbb{E}\,\eta$ oraz

$$\mathbb{E}\,(X;A) := \int_A X \mathrm{d}\mathbb{P} = \mathbb{E}\,(\eta;A), \qquad \forall A \in \mathcal{G}.$$

Warunkowa wartość oczekiwana wyznaczona jest jednoznaczne z dokładnością do zbioru $\mathbb{P}$-miary zero. Będziemy ją oznaczać przez $\mathbb{E}\,(X|\mathcal{G})$.

Istnienie warunkowej wartości oczekiwanej $\mathbb{E}\,(X|\mathcal{G})$ wynika z twierdzenia Radona–Nikodyma. Mianowicie

$$\mathbb{Q}(A) := \mathbb{E}\,(X;A), \qquad A \in \mathcal{G}$$

definiuje miarę σ-skończoną na $(\Omega, \mathcal{G}, \mathbb{P})$. Jest ona absolutnie ciągła ze względu na $\mathbb{P}$, to znaczy zachodzi

$$\mathbb{P}(A) = 0 \Longrightarrow \mathbb{Q}(A) = 0, \qquad \forall A \in \mathcal{G}.$$

Z twierdzenia Radona–Nikodyma wynika istnienie gęstości

$$\eta = \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}},$$

to znaczy $\mathcal{G}$ mierzalnej funkcji takiej, że

$$\mathbb{Q}(A) = \int_A \eta \mathrm{d}\mathbb{P}, \qquad \forall A \in \mathcal{G}.$$

Oczywiście gęstość ta ma żądane własności warunkowej wartości oczekiwanej.

Wypiszmy podstawowe własności warunkowej wartości oczekiwanej. Ich proste dowody zostawiamy czytelnikowi.

Twierdzenie 1.1 *(i) Jeżeli $X \geq 0$, $\mathbb{P}$-p.n. to $\mathbb{E}\,(X|\mathcal{G}) \geq 0$.*
(ii) $\mathbb{E}\,(1|\mathcal{G}) = 1$, $\mathbb{P}$-p.n.
(iii) Dla dowolnych zmiennych losowych X i Y, dla których $\mathbb{E}\,|X| < +\infty$ i $\mathbb{E}\,|Y| < +\infty$, $\mathbb{P}$-p.n. mamy

$$\mathbb{E}\,(X+Y|\mathcal{G}) = \mathbb{E}\,(X|\mathcal{G}) + \mathbb{E}\,(Y|\mathcal{G}).$$

(iv) Dla dowolnych zmiennych losowych X i Y, takich że Y jest $\mathcal{G}$-mierzalne i $\mathbb{E}\,|X| < +\infty$ i $\mathbb{E}\,|YX| < +\infty$, $\mathbb{P}$-p.n. mamy

$$\mathbb{E}\,(XY|\mathcal{G}) = Y\,\mathbb{E}\,(X|\mathcal{G}).$$

(v) Jeżeli $\mathcal{H}$ jest pod-σ-ciałem $\mathcal{G}$, to $\mathbb{P}$-p.n. mamy

$$\mathbb{E}\,(X|\mathcal{H}) = \mathbb{E}\,(\mathbb{E}\,(X|\mathcal{G})|\mathcal{H})\,.$$

(vi) Jeżeli X jest niezależne od $\mathcal{G}$, to znaczy jeżeli

$$\mathbb{P}\left(\{X \in \Gamma\} \cap A\right) = \mathbb{P}\left(X \in \Gamma\right)\, \mathbb{P}\left(A\right), \qquad \forall\, A \in \mathcal{G},$$

to

$$\mathbb{E}\left(X|\mathcal{G}\right) = \mathbb{E}\, X.$$

W szczególności, powyższa równość zachodzi gdy $\mathcal{G} = \{\emptyset, \Omega\}$.

Załóżmy, że Y jest $\mathcal{F}$-mierzalnym odwzorowaniem Ω w jakąś przestrzeń mierzalną $(E, \mathcal{E})$. Niech $\mathcal{Y} := \sigma(Y)$ będzie najmniejszym pod-σ-ciałem $\mathcal{F}$ względem którego Y jest mierzalne. Wówczas z definicji kładziemy

$$\mathbb{E}\left(X|Y\right) = \mathbb{E}\left(X|\sigma(Y)\right).$$

Mamy następujący fakt.

Lemat 1.1 *Rzeczywista zmienna losowa* η *mierzalna względem* σ*-ciała* $\sigma(Y)$ *jest postaci*

$$\eta = \psi(Y),$$

gdzie $\psi\colon E \to \mathbb{R}$ *jest odpowiednio dobraną funkcją mierzalną.*

Dowód Załóżmy, że η jest zmienną losową prostą, to znaczy przyjmującą skończoną liczbę wartości,

$$\eta = \sum_{i=1}^{M} \alpha_i \chi_{A_i}, \qquad A_i \cap A_j = \emptyset, \quad i \neq j, \;\; A_j \in \sigma(Y).$$

Zauważmy, że rodzina zbiorów

$$\{\omega \in \Omega\colon Y(\omega) \in \Gamma\}, \qquad \Gamma \in \mathcal{E}$$

tworzy σ-ciało i dlatego jest identyczna z $\sigma(Y)$. Stąd dla dowolnego i istnieje $\Gamma_i \in \mathcal{E}$, że

$$\chi_{A_i}(\omega) = \chi_{\Gamma_i}(Y(\omega)), \qquad \omega \in \Omega.$$

Czyli

$$\eta(\omega) = \sum_{i=1}^{M} \alpha_i \chi_{\Gamma_i}(Y(\omega)) = \psi(Y(\omega)), \qquad \omega \in \Omega,$$

gdzie

$$\psi(x) = \sum_{i=1}^{M} \alpha_i \chi_{\Gamma_i}(x), \qquad x \in E.$$

Niech η będzie dowolną zmienną losową. Wówczas istnieje ciąg zmiennych losowych prostych $\{\eta_n\}$, taki że

$$\eta(\omega) = \lim_{n\to\infty} \eta_n(\omega), \qquad \omega \in \Omega.$$

Wówczas $\eta_n = \psi_n(Y)$ i wystarczy przyjąć

$$\psi(x) = \limsup_{n\to\infty} \psi_n(x), \qquad x \in E.$$

Funkcja ψ jest mierzalna i $\eta = \psi(Y)$. $\square$

Z lematu wynika, że

$$\mathbb{E}\,(X|Y) = \psi(Y)$$

dla odpowiednio dobranej funkcji mierzalnej $\psi\colon E \to \mathbb{R}$. Funkcja ψ jest określona jednoznacznie ze względu na rozkład $\mathcal{L}(Y)$ elementu losowego Y. Kładziemy

$$\mathbb{E}\,(X|Y = y) := \psi(y)$$

dla $\mathcal{L}(Y)$ prawie wszystkich $y \in E$.

Następujący lemat pokazuje, że w przypadku zmiennej losowej przyjmującej skończoną liczbę wartości, pojęcie warunkowej wartości oczekiwanej jest identyczne z tradycyjnym pojęciem warunkowej wartości oczekiwanej ze względu na zdarzenie.

Lemat 1.2 *Niech $G_1, \ldots, G_M$ będą rozłącznymi elementami $\mathcal{F}$, takimi że*

$$\mathbb{P}\,(G_j) \neq 0, \quad j = 1, \ldots M \quad \text{i} \quad \bigcup_{i=1}^{M} G_j = \Omega.$$

Wówczas

$$\mathbb{E}\,(X|\mathcal{G})\chi_{G_j} = \frac{\mathbb{E}\,(Z; G_j)}{\mathbb{P}\,(G_j)}\,\chi_{G_j}, \qquad j = 1, \ldots, M.$$

Dowód Dowolna zmienna losowa mierzalna ze względu na

$$\mathcal{G} =: \sigma(G_1, \ldots, G_M)$$

jest stała na zbiorach G_j, $j = 1, \ldots, M$. Miech $\mathbb{E}\,(X|\mathcal{G}) = a_j$ na G_j. Wtedy

$$\mathbb{E}\,(X; G_j) = a_j\,\mathbb{E}\,\chi_{G_j} = a_j\,\mathbb{P}\,(G_j).$$

$\square$

Propozycja 1.1 *Załóżmy, że $\mathbb{E}\,X^2 < \infty$. Wtedy $\mathbb{E}\,(X|\mathcal{G})$ jest rzutem ortogonalnym X na podprzestrzeń $L^2(\Omega, \mathcal{G}, \mathbb{P})$ przestrzeni $L^2(\Omega, \mathcal{F}, \mathbb{P})$.*

Dowód Oznaczmy przez $\tilde{X}$ rzut ortogonalny zmiennej losowej X na przestrzeń $L^2(\Omega, \mathcal{G}, \mathbb{P})$. Oczywiście iloczyn skalarny na $L^2(\Omega, \mathcal{G}, \mathbb{P})$ dany jest wzorem

$$\langle X_1, X_2 \rangle = \mathbb{E}\,X_1 X_2.$$

Wtedy

$$\left(X - \tilde{X}\right) \perp L^2(\Omega, \mathcal{G}, \mathbb{P}).$$

W szczególności

$$\mathbb{E}\,(X - \tilde{X})\chi_G = 0 \qquad \forall\, G \in \mathcal{G}.$$

Stąd wynika, że $\tilde{X} = \mathbb{E}\,(X|\mathcal{G})$. $\square$

1.3 Prawdopodobieństwo warunkowe

Niech $(E, \mathcal{E})$ będzie przestrzenią mierzalną. Oznaczmy przez $\mathcal{P}(E)$ przestrzeń wszystkich miar probabilistycznych na $(E, \mathcal{E})$. Na $\mathcal{P}(E)$ rozważamy najmniejsze σ-ciało $\mathfrak{P}(E)$, dla którego wszystkie odwzorowania

$$\mathcal{P}(E) \ni \mu \to \mu(A) \in \mathbb{R}, \qquad A \in \mathcal{E}$$

są mierzalne.

Niech $(\Omega, \mathcal{F}, \mathbb{P})$ będzie przestrzenią probabilistyczną, a $\mathcal{Y}$ pod-σ-ciałem $\mathcal{F}$.

Definicja 1.2 Niech X będzie elementem losowym w E określonym na $(\Omega, \mathcal{F}, \mathbb{P})$. *Rozkładem warunkowym* X względem $\mathcal{Y}$ nazywamy element losowy $\mathbb{P}(X|\mathcal{Y})(\omega)$, $\omega \in \Omega$, o wartościach w $(\mathcal{P}(E), \mathfrak{P}(E))$, taki że dla dowolnej funkcji mierzalnej ograniczonej ψ z prawdopodobieństwem 1 mamy

$$\mathbb{E}(\psi(X)|\mathcal{Y})(\omega) = \int_E \psi(x)\mathbb{P}(X|\mathcal{Y})(\omega)(\mathrm{d}x).$$

Oczywiście z prawdopodobieństwem 1 mamy

$$\mathbb{P}(X|\mathcal{Y})(A) = \mathbb{E}(\chi_A(X)|\mathcal{Y}), \qquad \forall\, A \in \mathcal{E}.$$

Formuła to definiuje rozkład warunkowy przy założeniu, że warunkowe wartości oczekiwane

$$\mathbb{E}(\chi_A(X)|\mathcal{Y})(\omega), \qquad \forall\, A \in \mathcal{E}$$

można określić w taki sposób, aby dla każdego $\omega \in \Omega$ zadawały miarę probabilistyczną na $\mathcal{F}$. Prawdopodobieństwo warunkowe istnieje, patrz [12], str. 13, gdy przestrzeń $(\Omega, \mathcal{F})$ jest dostatecznie regularna, na przykład gdy jest *standartową przestrzenią mierzalną*, to znaczy gdy jest izomorficzną z jedną z następujących przestrzeni

$$(\langle 1, n\rangle, \mathcal{B}(\langle 1, n\rangle)), \qquad (\mathbb{N}, \mathcal{B}(\mathbb{N}), \qquad (\mathbb{M}, \mathcal{B}(\mathbb{M}),$$

gdzie $\langle 1, n\rangle = \{1, \ldots, n\}$ z topologią dyskretną, $\mathbb{N}$ zbiór liczb naturalnych z topologią dyskretną, a

$$\mathbb{M} = \{0, 1\}^{\mathbb{N}}$$

z topologią produktową. Wiadomo, patrz [13], [16], że przestrzeniami standartowymi są: dowolny mierzalny podzbiór przestrzeni standartowej z indukowanym σ-ciałem i dowolna *przestrzeń polska*, to jest zupełna ośrodkowa przestrzeń metryczna, z topologicznym σ-ciałem.

Definicja 1.3 Załóżmy że $(E_1, \mathcal{E}_1)$ i $(E_2, \mathcal{E}_2)$ są przestrzeniami mierzalnymi. Niech (X, Y) będzie wektorem losowym w $E_1 \times E_2$ określonym na przestrzeni probabilistycznej $(\Omega, \mathcal{F}, \mathbb{P})$, oraz niech $\mathcal{Y}$ będzie σ-ciałem generowanym przez Y. Wówczas rozkład warunkowy X względem $\mathcal{Y}$ nazywamy *rozkładem warunkowym X względem Y*.

Dla zadanego wektora $m \in \mathbb{R}^d$ i symetrycznej dodatnio określonej macierzy $Q \in M(d \times d)$ niech $N(m, Q)$ oznacza rozkład gaussowski o średniej m i macierzy kowariancji Q. Gdy Q jest ściśle dodatnio określona to rozkład $\mathcal{N}(m, Q)$ ma gęstość

$$g(x) = \frac{1}{\sqrt{(2\pi)^n \det Q}} \exp\left\{-\frac{1}{2}\left\langle Q^{-1}(x-m), x-m\right\rangle\right\}, \qquad x \in \mathbb{R}^d.$$

Definicja 1.4 Rozkład warunkowy X względem $\mathcal{Y}$ jest *gaussowski* gdy

$$\mathbb{P}(X|\mathcal{Y})(\omega) = N(m(\omega), Q(\omega)), \qquad \omega \in \Omega,$$

gdzie m i Q to zmienne losowe o wartościach odpowiednio wektorowych i macierzowych.

Poniższy przykład powinien ułatwić przyswojenie pojęcia warunkowego rozkładu oczekiwanego. Mianowicie, załóżmy, że X i Y są zmiennymi losowymi o wartościach w przestrzeniach mierzalnych $(E_1, \mathcal{E}_1)$ i $(E_2, \mathcal{E}_2)$. Załóżmy, że istnieje gęstość $g(x, y)$, $(x, y) \in E_1 \times E_2$ rozkładu (X, Y) względem miary probabilistycznej $\lambda_1 \otimes \lambda_1$ na $E_1 \times E_2$, to znaczy, że

$$\mathbb{P}((X, Y) \in \Gamma) = \int_\Gamma g(x, y)\lambda_1(\mathrm{d}x)\lambda_2(\mathrm{d}y), \qquad \Gamma \in \mathcal{E}_1 \otimes \mathcal{E}_2.$$

Zdefiniujmy

$$g(x|y) = \begin{cases} \frac{g(x,y)}{\int_{E_1} g(z,y)\lambda_1(\mathrm{d}z)} & \text{gdy } \int_{E_1} g(z, y)\lambda_1(\mathrm{d}z) \neq 0, \\ g_1(x) & \text{gdy } \int_{E_1} g(z, y)\lambda_1(\mathrm{d}z) \neq 0, \end{cases}$$

gdzie $g_1 \geq 0$ jest dowolną funkcją spełniającą

$$\int_{E_1} g_1(z)\lambda_1(\mathrm{d}z) = 0.$$

Funkcję g nazywamy *gęstością warunkową*.

Lemat 1.3 *Zachodzi*

$$\mathbb{P}(X|Y)(\omega)(\mathrm{d}x) = g(x|Y(\omega))\lambda_1(\mathrm{d}x).$$

Dowód Należy pokazać, że dla dowolnej funkcji $\psi \colon E_1 \to \mathbb{R}_+$ mamy

$$\mathbb{E}(\psi(X)|\sigma(Y)) = \int_{E_1} \psi(x)g(x|Y)\lambda_1(\mathrm{d}x).$$

Niech

$$\tilde{Z} = \int_{E_1} \psi(x)g(x|Y)\lambda_1(\mathrm{d}x)$$

oraz niech $\Gamma \in \mathcal{E}_2$. Gęstość zmiennej Y względem λ_2 ma postać

$$\int_{E_1} g(x,y)\lambda_1(\mathrm{d}x), \qquad y \in E_2.$$

Dlatego

$$\begin{aligned}
&\mathbb{E}\,\tilde{Z}\chi_{\Gamma}(Y)\\
&= \mathbb{E}\left(\int_{E_1} \psi(x)g(x|Y)\lambda_1(\mathrm{d}x)\,\chi_{\Gamma}\right)\\
&= \int_{E_2} \chi_{\Gamma}\left(\int_{E_1} \psi(x)g(x|Y)\lambda_1(\mathrm{d}x)\right)\left(\int_{E_1} g(z,y)\lambda_1(\mathrm{d}z)\right)\lambda_2(\mathrm{d}y)\\
&= \int_{E_2} \chi_{\Gamma}(y)\int_{E_1} \psi(x)g(x,y)\lambda_2(\mathrm{d}y)\\
&= \mathbb{E}\,\psi(X)\chi_{\Gamma}(Y).
\end{aligned}$$

□

Rozdział zakończymy następującym rezultatem.

Lemat 1.4 *Niech $\mathbb{P}(X|\mathcal{Y})$ będzie rozkładem warunkowym elementu losowego X o wartościach w $(E_1,\mathcal{E}_1)$ względem σ ciała $\mathcal{Y}$. Niech Y będzie $\mathcal{Y}$ mierzalnym elementem losowym o wartościach w przestrzeni mierzalnej $(E_2,\mathcal{E}_2)$. Wówczas dla dowolnej funkcji mierzalnej ograniczonej $\psi\colon E_1 \times E_2 \to \mathbb{R}$ z prawdopodobieństwem 1 zachodzi*

$$\mathbb{E}\,(\psi(X,Y)|\mathcal{Y}) = \int_{E_1} \psi(x,Y)\mathbb{P}\,(X|\mathcal{Y})(\mathrm{d}x).$$

Dowód Teza jest prawdziwa dla funkcji ψ postaci

$$\psi(x,y) = \chi_{\Gamma_1}(x)\chi_{\Gamma_2}(y),$$

gdzie $\Gamma_i \in \mathcal{E}_i$, $i = 1,2$. Rodzina zbiorów mierzalnych $\Gamma \subset E_1 \times E_2$ dla których

$$\mathbb{E}\,(\chi_{\Gamma}(X,Y)|\mathcal{Y}) = \int_{E_1} \chi_{\Gamma}(x,Y)\mathbb{P}\,(X|\mathcal{Y})(\mathrm{d}x)$$

jest π-układemem, patrz [10]. Stąd równa jest całej σ-algebrze $\mathcal{E}_1 \times \mathcal{E}_2$. Mając równość dla funkcji charakterystycznych, przez przejście graniczne łatwo otrzymujemy ją dla dowolnej funkcji mierzalnej ograniczonej. □

1.4 Filtracje i momenty Markowa

Niech $(\Omega,\mathcal{F},\mathbb{P})$ będzie przestrzenią probabilistyczną. *Filtracją* nazywamy dowolny niemalejący ciąg $(\mathcal{F}_n)$ pod-σ-ciał $\mathcal{F}$, to znaczy ciąg spełniający

$$\mathcal{F}_n \subseteq \mathcal{F}_m, \qquad \forall\, n \le m.$$

Jeżeli (X_n) jest ciągiem elementów losowych na $(\Omega, \mathcal{F}, \mathbb{P})$ przyjmujących wartości w być może różnych przestrzeniach mierzalnych $(E_n, \mathcal{E}_n)$, to *filtracją generowaną przez* (X_n) jest ciąg $(\mathcal{X}_n)$,

$$\mathcal{X}_n := \sigma(X_1, X_2, \ldots, X_n), \qquad n = 1, 2, \ldots,$$

gdzie $\sigma(X_1, \ldots, X_n)$ oznacza najmniejsze pod-σ-ciało $\mathcal{F}$ względem, którego wszystkie elementy $X_1, \ldots, X_n$ są mierzalne.

Definicja 1.5 Niech $(\Omega, \mathcal{F}, \mathbb{P})$ będzie przestrzenią probabilistyczną z filtracją $(\mathcal{F}_n)$. Zmienną losową

$$\tau \colon \Omega \to 1, 2, \ldots \cup \{+\infty\}$$

nazywamy *momentem Markowa względem filtracji* $(\mathcal{F}_n)$ gdy

$$\{\tau \le n\} \in \mathcal{F}_n, \qquad \forall\, n.$$

Łatwo wykazać, że jeżeli τ jest momentem Markowa względem $(\mathcal{F}_n)$, to

$$\{\tau = n\},\ \{\tau < n\},\ \{\tau \ge n\},\ \{\tau > n\}\ \in \mathcal{F}_n, \qquad \forall\, n.$$

1.5 Łańcuch Markowa

Zasadniczym tematem wykładów jest sterowanie *łańcuchem Markowa*, inaczej *procesem Markowa w czasie dyskretnym.* Rozważmy przestrzeń probabilistyczną $(\Omega, \mathcal{F}, \mathbb{P})$ z filtracją $(\mathcal{F}_n)$. Niech E będzie przestrzenią mierzalną. Będziemy interpretowali E jako *przestrzeń stanów* naszego procesu. Proces (X_n) o wartościach w E nazywamy $(\mathcal{F}_n)$-*adaptowany* gdy dla dowolnego n, X_n jest $\mathcal{F}_n$-mierzalny. Oznaczmy przez $B_b(E)$ przestrzeń wszystkich funkcji rzeczywistych określonych na E, mierzalnych i ograniczonych.

Definicja 1.6 Niech będzie $(\mathcal{F}_n)$-adaptowanym procesem w wartościach w E oraz niech (P_n) będzie ciągiem operatorów liniowych z $B_b(E)$ w $B_b(E)$. Wówczas $(X_n, \mathcal{F}_n)$ jest *łańcuchem Markowa z rodziną operatorów przejścia* (P_n) gdy

$$\mathbb{E}\left(\psi(X_{n+1}) | \mathcal{F}_n\right) = P_n \psi(X_n), \qquad \forall\, n = 0, 1, \ldots,\ \forall\, \psi \in B_b(E).$$

Łańcuch Markowa (X_n) jest *jednorodny* gdy rodzina operatorów przejścia nie zależy od n, to znaczy gdy istnieje operator liniowy na $B_b(E)$, taki że

$$\mathbb{E}\left(\psi(X_{n+1}) | \mathcal{F}_n\right) = P \psi(X_n), \qquad \forall\, n = 0, 1, \ldots,\ \forall\, \psi \in B_b(E).$$

Podstawowym obiektem naszych dalszych rozważań będzie sterowany łańcuch Markowa.

Definicja 1.7 Niech $(\Omega, \mathcal{F}, \mathbb{P})$ będzie przestrzeńą probabilistyczną z filtracją $(\mathcal{F}_n)$. Niech $(E, \mathcal{E})$ będzie mierzalną przestrzenią stanów a $(U,\mathcal{U})$ mierzalną przestrzenią, którą interpretujemy jako *przestrzeń sterowań*. Niech $N < \infty$ będzie skończonym horyzontem czasowym oraz niech

$$P_n^u \colon B_b(E) \to B_b(E), \qquad n = 0, 1, \ldots, N-1,\ u \in U$$

będzie rodziną operatorów liniowych na przestrzeni funkcji $B_b(E)$ mierzalnych ograniczonych na E. Zakładamy, że dla dowolnych $\psi \in B_b(E)$ i $x \in E$, odwzorowanie

$$U \ni u \to P^u\psi(x) \in E$$

jest mierzalne. *Strategią* nazywamy dowolny $(\mathcal{F}_n)$-adaptowany proces

$$\pi = (u_0, u_1, \ldots, u_{n-1})$$

o wartościach w U. Załóżmy, że dla dowolnej strategii π określony jest $(\mathcal{F}_n)$-adaptowany proces $X = (X_n^\pi)$ o wartościach w E, taki że

$$\mathbb{E}\left(\psi(X_{n+1}^\pi)|\mathcal{F}_n\right) = P_n^{u_n}\psi(X_n^\pi), \qquad n = 0, 1, \ldots N-1,\ \psi \in B_b(E).$$

Proces $(X_n^\pi, \mathcal{F}_n)$ nazywamy *sterowanym łańcuchem Markowa.*

1.6 Postać rekurencyjna

Załóżmy, że proces (X_n) zadany jest rekurencyjnie

$$X_{n+1} = F_n(X_n, \xi_{n+1}), \quad n = 0, 1, \ldots,$$

gdzie $\{\xi_n\}$ jest ciągiem niezależnych zmiennych losowych określonych na przestrzeni probabilistycznej $(\Omega, \mathcal{F}, \mathbb{P})$, przyjmujących wartości w przestrzeni mierzalnej $(S, \mathcal{S})$, a (F_n) jest ciągiem mierzalnych odwzorowań produktu $E \times S$ w E. Ponadto, zakładamy, że znamy stan początkowy X_0, który jest zmienną losową w E niezależną od ciągu $\{\xi_n\}$. Niech

$$\mathcal{X}_n = \sigma(X_0, \ldots, X_n), \qquad n = 0, 1, \ldots.$$

Mamy następujący rezultat.

Twierdzenie 1.2 *Proces $(X_n, \mathcal{X}_n)$ jest łańcuchem Markowa z ciągiem operatorów przejścia*

$$P_n\psi(x) = \int_S \psi(F_n(x, w))\mathcal{L}(\xi_{n+1})(\mathrm{d}w), \qquad \psi \in B_b(E),\ n = 0, 1, \ldots,$$

gdzie $\mathcal{L}(\xi_{n+1})$ oznacza rozkład ξ_{n+1} w $(S, \mathcal{S})$.

Dowód Ustalmy ψ i n. Ponieważ ξ_{n+1} jest niezależne od $\mathcal{X}_n$ i X_n jest $\mathcal{X}_n$-mierzalne mamy

$$\mathbb{P}\left(\xi_{n+1}|\mathcal{X}_n\right) = \mathcal{L}\left(\xi_{n+1}\right).$$

Z Lematu 1.4,

$$\begin{aligned}\mathbb{E}\left(\psi(X_{n+1})|\mathcal{X}_n\right) &= \mathbb{E}\left(\psi(F(X_n,\xi_{n+1}))|\mathcal{X}_n\right)\\ &= \int_S F(X_n,w)\mathcal{L}(\xi_{n+1})(\mathrm{d}w),\end{aligned}$$

co kończy dowód. □

Oczywiście gdy zmienne (ξ_n) mają ten sam rozkład, oznaczmy go przez $\mathcal{L}(\xi)$ i gdy $F_n = F_m =: F$ dla wszystkich m, n to proces $(X_n, \mathcal{X}_n)$ jest jednorodny z operatorem przejścia

$$P\psi(x) = \int_S \psi(F(x,w))\mathcal{L}(\xi)(\mathrm{d}w), \qquad \psi \in B_b(E).$$

Analogiczne twierdzenie, z analogicznym dowodem, zachodzi dla sterowanego łańcucha Markowa.

Załóżmy, że dane są przestrzenie mierzalne $(E,\mathcal{E})$, $(U,\mathcal{U})$ i $(S,\mathcal{S})$. Niech (ξ_n) będzie ciągiem niezależnych zmiennych losowych o wartościach w S. Niech

$$\mathcal{X}_n = \sigma(\xi_0,\ldots,\xi_n).$$

Dla strategii

$$\pi = (u_0, u_n, \ldots),$$

to jest $(\mathcal{X}_n)$-adoptowalnego procesu definiujemy proces (X_n^π) rekurencyjnie

$$X_{n+1} = F_n(X_n, u_n, \xi_{n+1}), \quad n = 0,1,\ldots.$$

Twierdzenie 1.3 $(X_n^\pi, \mathcal{X}_n)$ *jest sterowanym łańcuchem Markowa z operatorami przejścia*

$$P_n^u\psi(x) = \int_S F(x,u,w)\mathcal{L}(\xi_{n+1})(\mathrm{d}w), \qquad n = 0,1,\ldots,\ \psi \in B_b(E).$$

Pokazaliśmy jak zdefiniowć łańcuch Markowa. Następujący fakt pokazuje, że każdy łańcuch Markowa $(X_n, \mathcal{X}_n)$ daje się zdefiniować w postaci rekurencyjnej. Dla prostoty ograniczymy się do jednorodnych łańcuchów Markowa.

Twierdzenie 1.4 *Niech E będzie przestrzenią polską, $\mathcal{E}$ niech będzie σ-algebrą zbiorów borelowskich na E, a $(U,\mathcal{U})$ niech będzie dowolną przestrzenią mierzalną. Niech $P^u(x,\Gamma)$, $x \in E$, $u \in U$, $\Gamma \in \mathcal{E}$, będzie prawdopodobieństwem przejścia na E, to znaczy, zakładamy, że:*

a) dla dowolnych $x \in E$ oraz $u \in U$, $P^u(x,\cdot)$ jest miarą probabilistyczną na $(E,\mathcal{E})$, *b) dla dowolnego* $\Gamma \in \mathcal{E}$, *odwzorowanie* $(x,u) \to P^u(x,\Gamma)$ *jest mierzalne z* $E \times U$ *w* $[0,1]$.

Niech μ *będzie miarą probabilistyczną na* E. *Istnieją wówczas odwzorowania mierzalne*

$$F\colon E \times U \times [0,1] \to E, \qquad f\colon [0,1] \to E,$$

takie że dla ciągu niezależnych zmiennych losowych $\xi_0, \xi_1, \ldots,$ *o rozkładzie jednostajnym na* $[0,1]$, *proces*

$$X_{n+1} = F(X_n, u_n, \xi_{n+1}), \qquad n = 0, 1, \ldots$$

ma operator przejścia P^u, *a* $X_0 = f(\xi_0)$ *ma rozkład* μ.

2

Sterowanie na przedziale skończonym

2.1 Intuicyjne wprowadzenie do zasady indukcji wstecz

Załóżmy, że w każdym, momencie gracz może postawić dowolną część u swojego kapitału x i albo przegrać albo wygrać u odpowiednio z prawdopodobieństwem p i $q = 1 - p$. Załóżmy, że gracz ma prawo do N gier. Jego celem jest zmaksymalizowanie satysfakcji, którą jest logarytm końcowego kapitału. Niech $V_n(x)$ oznacza maksymalną średnią satysfakcję przy założeniu, że gracz zagrał już n razy a x jest kapitałem gracza w n-tej grze. Oczywiście najbardziej interesuje nas $V_0(x)$. Mamy

$$V_N(x) = \log x, \qquad x \geq 0.$$

Ponadto

$$\begin{aligned} V_{N-1}(x) &= \sup_{u\in[0,1]} \mathbb{E}\left(V(X_N^u)|X_{N-1} = x\right) \\ &= \sup_{u\in[0,1]} \left(pV(x + ux) + (1-p)V(x - ux)\right), \end{aligned}$$

gdzie X_N^u jest kapitałem gracza w chwili N przy założeniu, że w chwili $N - 1$ gracz obstawił u-tą część swojego kapitału. Tutaj korzystamy z faktu, że

$$X_N^u = x + \xi ux,$$

gdzie ξ jest zmienną losową przyjmującą wartość 1 z prawdopodobieństwem p i -1 z prawdopodobieństwem $q = 1 - p$. Tak więc z Lematu 1.4,

$$\mathbb{E}\left(V(X_N^u)|X_{N-1} = x\right) = \mathbb{E}\left(V(x + \xi ux)\right) = pV(x + ux) + (1-p)V(x - ux).$$

Ogólnie

$$V_{n-1}(x) = \sup_{u\in[0,1]} \mathbb{E}\left(V_n(X_n^u)|X_{n-1} = x\right)$$

gdzie X_n^u jest kapitałem gracza w chwili n jeśli w chwili $n - 1$ obstawił u-tą część swojego kapitału.

Mamy więc ciąg zadany indukcyjnie

$$\begin{aligned} V_N(x) &= \log x, \\ V_{n-1}(x) &= \sup_{u\in[0,1]} \mathbb{E}\left(V_n(X_n^u)|X_{n-1}=x\right) \\ &= \sup_{u\in[0,1]} \left(pV_n(x+ux)+qV_n(x-ux)\right). \end{aligned} \tag{2.1}$$

Co więcej w chwili n optymalne jest obstawienie części u kapitału, gdzie $u = u(x)$, przy zadanym x realizuje supremum

$$\sup_{u\in[0,1]} \left(pV_n(x+ux)+qV_n(x-ux)\right).$$

Inaczej

$$u(x) \in \operatorname{Arg}\sup_{u\in[0,1]} \left(pV_n(x+ux)+qV_n(x-ux)\right).$$

Pokażemy, że jeżeli $p \leq 1/2$ to $V_n(x) = \log x$, to znaczy optymalnie jest nic nie stawiać. Załóżmy, że $0 < m \leq N$ jest takie, że $V_N(x) = \ldots V_m(x) = \log x$ dla $x > 0$. Mamy pokazać, że $V_{m-1}(x) = \log x$. Istotnie

$$\begin{aligned} V_{m-1}(x) &= \sup_{u\in[0,1]} \left(p\log(x+ux)+(1-p)\log(x-ux)\right) \\ &= \log x + \sup_{u\in[0,1]} \left(p\log(1+u)+(1-p)\log(1-u)\right) \end{aligned}$$

Wystarczy więc zauważyć, że dla $p \leq 1/2$,

$$\sup_{u\in[0,1]} \left(p\log(1+u)+(1-p)\log(1-u)\right) = p\log 1 + (1-p)\log 1 = 0.$$

Rozważamy teraz przypadek $p > 1/2$. Pokażemy, że w tym przypadku

$$V_{N-n}(x) = \log x + nC,$$

gdzie

$$C = \sup_{u\in[0,1]} \left(p\log(1+u)+(1-p)\log(1-u)\right).$$

Ponadto supremum jest osiągalne (niezależnie od n i x) w punkcie

$$\widehat{v}(x) = 2p-1 = p-q.$$

Po pierwsze zauważmy, że rzeczywiście supremum po prawej stronie równości na C jest osiągalne dla $u = p - q$. Następnie zauważmy, że dla $n = 0$ równość zachodzi z definicji V_N. Jeżeli zachodzi dla n, to

$$\begin{aligned} V_{N-n-1}(x) &= \sup_{u\in[0,1]} \left[pV_{N-n}\left(x+ux\right)+(1-p)V_{N-n}\left(x-ux\right)\right] \\ &= nC + \sup_{u\in[0,1]} \left[p\log\left(x+ux\right)+(1-p)\log\left(x-ux\right)\right] \\ &= nC + \log x + \sup_{u\in[0,1]} \left(p\log(1+u)+(1-p)\log(1-u)\right) \\ &= (n+1)C + \log x. \end{aligned}$$

2.2 Formalne postawienie problemu

Na początku rozważamy przypadek sterowania procesem zdefiniowanym rekurencyjnie. Dla prostoty oznaczeń zakładamy, że proces jest jednorodny w czasie. Tak więc

$$X_{n+1} = F(X_n, u_n, \xi_{n+1}), \quad n = 0, 1, \ldots, \tag{2.2}$$

gdzie $\{\xi_n\}$ jest ciągiem iid zmiennych losowych określonych na przestrzeni probabilistycznej $(\Omega, \mathcal{F}, \mathbb{P})$, przyjmujących wartości w przestrzeni mierzalnej $(S, \mathcal{S})$, sterowania $\{u_n\}$ przyjmują wartości w przestrzeni mierzalnej $(U, \mathcal{U})$, $\{X_n\}$ przyjmuje wartości w mierzalnej przestrzeni stanów $(E, \mathcal{E})$, a F jest mierzalnym odwzorowaniem produktu $E \times U \times S$ w E. Zakładamy ponadto, że dla dowolnego $u \in E$ zadany jest zbiór $U(x) \in \mathcal{U}$. Interpretujemy go jako zbiór sterowań dopuszczalnych przy warunku, że proces jest w stanie x. Ponadto, zakładamy, że znamy stan początkowy X_0, który jest zmienną losową w E niezależną od ciągu $\{\xi_n\}$.

W Przykładzie 1.1, $X_n \geq 0$, a więc $E = [0, +\infty)$, $\xi_n \in [-1, +\infty)$, a więc $S = [-1, +\infty)$, $u_n = (c_n, b_n) \in [0, x_n] \times [0, 1]$. Stąd $U = [0, \infty) \times [0, 1]$ oraz $U(x) = [0, x] \times [0, 1]$. Ponadto

$$\begin{aligned} X_{n+1} &= b_n(X_n - c_n)(1 + r) + (1 - b_n)(1 + \xi_{n+1})(X_n - c_n) \\ &= (X_n - c_n)\left(b_n + b_n r + 1 + \xi_{n+1} - b_n - b_n \xi_{n+1}\right) \\ &= (X_n - c_n)\left(1 + \xi_{n+1} + b_n(r - \xi_{n+1})\right). \end{aligned}$$

Tak więc

$$F(x, (c, b), \xi) = (x - c)\left(1 + \xi + b(r - \xi)\right).$$

Definicja 2.1 *Strategią* nazywamy dowolny ciąg $\pi = (u_n)$ odwzorowań mierzalnych

$$u_n \colon E^{n+1} \to U,$$

taki, że

$$u_n(x_0, \ldots, x_n) \in U(x_n), \qquad \forall\, n,\ \forall\, (x_0, \ldots, x_n) \in E^{n+1}.$$

Proces X^π odpowiadający strategii π, to jest *odpowiedz na* π, definiujemy w następujący sposób

$$X^\pi_{n+1} = F(X^\pi_n, u_n(X^\pi_0, \ldots, x^\pi_n), \xi_{n+1}), \qquad X^\pi_0 = X_0.$$

W zapisie będziemy często opuszczali indeks π. Celem będzie znalezienie strategii, która maksymalizuje zysk lub minimalizuje koszt. Będziemy rozróżniali dwa przypadki.

I) *Sterowanie ze skończonym horyzontem czasowym* N. Wówczas zadany jest skończony przedział czasowy $[0, 1, \ldots, N-1, N]$ oraz zysk (lub koszt) postaci

$$\sum_{n=0}^{N-1} q_n(X_n^\pi, u_n(X_0^\pi, \dots, X_n^\pi)) + r_N(X_N^\pi).$$

Zysk (koszt) zależy od strategii π i wartości początkowej X_0. Celem jest maksymalizacja (odp. minimalizacja) funkcjonału (średniego) zysku (odp. kosztu)

$$J_N(\pi, X_0) = \mathbb{E}\left\{\sum_{n=0}^{N-1} q_n(X_n^\pi, u_n(X_0^\pi, \dots, X_n^\pi)) + r_N(X_N^\pi)\right\}. \tag{2.3}$$

Będziemy zakładali, że $q_n\colon E\times U \to \mathbb{R}$ oraz $r_N\colon E \to \mathbb{R}$ są zadanymi funkcjami mierzalnymi. Dla uproszczenia zakładać będziemy, że wszystkie funkcje $\{q_n\}$ oraz funkcja r_N są nieujemne.

Dla problemu inwestora, Przykład 1.1, J_N dany (1.1) jest funkcjonałem zysku.

II) *Sterowanie z nieskończonym horyzontem czasowym.* Tutaj $N = +\infty$. Zadany jest ciąg funkcji nieujemnych mierzalnych $\{q_n\}$. Funkcjonał kosztu (zysku) ma postać

$$J_\infty(\pi, X_0) = \mathbb{E}\sum_{n=0}^{+\infty} q_n(X_n^\pi, u_n(X_0^\pi, \dots, X_n^\pi)).$$

Oczywiście dla funkcjonału zysku szukamy strategii, która go maksymalizuje, a dla funkcjonału kosztu, staramy się go zminimalizować.

2.3 Zasada indukcji wstecz Bellmana

Ogólne twierdzenie dotyczące optymalnego sterowania na przedziale skończonym pochodzi od Richarda Bellmana. Oznaczmy przez P^u, $u \in U$ operatory przejścia dla procesu sterowanego. Z Twierdzenia 1.2 mamy

$$P^u h(x) = \mathbb{E}\, h(F(x, u, \xi_0)) = \mathbb{E}\left(h(X_{n+1}) | X_n = x, u_n = u\right). \tag{2.4}$$

Z definicji operatora przejścia mamy następujący lemat.

Lemat 2.1 *Zachodzi wzór* $\mathbb{E}\, P^u h(X_n) = \mathbb{E}\, h(X_{n+1})$.

Twierdzenie 2.1 (Zasada indukcji wstecz Bellmana) *Załóżmy, że dany jest łańcuch* (2.2)*, skończony horyzont czasowy N, oraz funkcjonał zysku (kosztu) postaci* (2.3)*. Niech V_N, V_{N-1}, …, V_0 będą mierzalnymi funkcjami na E, takimi że, jeżeli J_N jest funkcjonałem zysku to*

$$V_N(x) = r_N(x), \qquad V_n(x) = \sup_{u\in U(x)} \left(q_n(x, u) + P^u V_{n+1}(x)\right), \tag{2.5}$$

a jeżeli J_N jest funkcjonałem kosztu,

$$V_N(x) = r_N(x), \qquad V_n(x) = \inf_{u \in U(x)} \left(q_n(x,u) + P^u V_{n+1}(x) \right). \tag{2.6}$$

Wówczas dla dowolnej strategii π, $J_N(X_0, \pi) \leq \mathbb{E}\, V_0(X_0)$, gdy J_N jest zyskiem, a $J_N(X_0, \pi) \geq \mathbb{E}\, V_0(X_0)$, gdy J_N jest kosztem. Ponadto, jeżeli istnieją odwzorowania mierzalne $\widehat{v}_0, \ldots, \widehat{v}_{N-1} \colon E \to U$ takie, że dla wszystkich n i $x \in E$,

$$\widehat{v}_n(x) \in U(x), \qquad V_n(x) = q_n(x, \widehat{v}_n(x)) + P^{\widehat{v}_n(x)} V_{n+1}(x), \tag{2.7}$$

to strategia $\widehat{\pi} = \{\widehat{u}_n\}$, gdzie

$$\widehat{u}_n(x_0, x_1, \ldots, x_n) = \widehat{v}_n(x_n) \tag{2.8}$$

jest optymalna, oraz $J_N(\widehat{\pi}, X_0) = \mathbb{E}\, V_0(X_0)$.

Dowód Załóżmy, że J_N jest funkcjonałem zysku. Niech $\{V_n\}$ spełniają (2.5). Wówczas dla dowolnych $n = 0, \ldots, N_1$, $x \in E$ oraz $u \in U(x)$ mamy

$$V_n(x) \geq q_n(x,u) + P^u V_{n+1}(x).$$

Niech π będzie strategią i niech $X = X^\pi$. Wówczas z Lematu 2.1,

$$\begin{aligned} \mathbb{E}\, V_n(X_n) &\geq \mathbb{E}\, q_n(X_n, u_n(X_0, \ldots, X_n)) + \mathbb{E}\, P^{u_n(X_0, \ldots, X_n)} V_{n+1}(X_n) \\ &\geq \mathbb{E}\, q_n(X_n, u_n(X_0, \ldots, X_n)) + \mathbb{E}\, V_{n+1}(X_{n+1}). \end{aligned}$$

Sumując stronami i korzystając z postaci J_N otrzymujemy

$$\begin{aligned} \sum_{n=0}^{N-1} \mathbb{E}\, V_n(X_n) &\geq \sum_{n=0}^{N-1} \mathbb{E}\, q_n(X_n, u_n(X_0, \ldots, X_n)) + \sum_{n=0}^{N-1} \mathbb{E}\, V_{n+1}(X_{n+1}) \\ &\geq J_N(\pi, X_0) + \sum_{n=1}^{N-1} \mathbb{E}\, V_n(X_n). \end{aligned}$$

Jeżeli dla wszystkich n, $\mathbb{E}\, V_n(X_n) < +\infty$ to odejmując od obu stron wyrażenie

$$\sum_{n=1}^{N-1} \mathbb{E}\, V_n(X_n)$$

otrzymujemy żądane oszacowanie

$$\mathbb{E}\, V_0(X_0) \geq J_N(\pi, X_0). \tag{2.9}$$

Jeżeli dla jakiegoś $n > 0$ zachodzi $\mathbb{E}\, V_n(X_n) = +\infty$, to ponieważ

$$\mathbb{E}\, V_{n-1}(X_{n-1}) \geq \mathbb{E}\, q_n(X_n, U_n(X_0, \ldots, X_n)) + \mathbb{E}\, V_n(X_n)) = +\infty,$$

zachodzi również $\mathbb{E}\, V_{n-1}(X_{n-1}) = +\infty$, a w konsekwencji $\mathbb{E}\, V_0(X_0) = +\infty$. W tym przypadku, oczywiście zachodzi (2.9). Stosując to samo rozumowanie dla funkcjonału kosztu otrzymujemy

$$\sum_{n=0}^{N-1} \mathbb{E}\, V_n(X_n) \le J_N(\pi, X_0) + \sum_{n=1}^{N-1} \mathbb{E}\, V_n(X_n).$$

Tak więc jeżeli $\mathbb{E}\, V_n(X_n) < +\infty$ dla wszystkich n, to odejmując stronami otrzymujemy

$$\mathbb{E}\, V_0(X_0) \le J_N(\pi, X_0). \tag{2.10}$$

Jeżeli dla pewnego n, $\mathbb{E}\, V_n(X_n) = +\infty$ to

$$\mathbb{E}\, V_n(X_n) \le \mathbb{E}\, q_n(X_n, U_n(X_0, \ldots, X_n)) + \mathbb{E}\, V_{n+1}(X_{n+1}).$$

Stąd

$$\mathbb{E}\, q_n(X_n, U_n(X_0, \ldots, X_n)) = +\infty,$$

a więc

$$J_N(\pi, X_0) = +\infty \quad \text{lub} \quad \mathbb{E}\, V_{n+1}(X_{n+1}) = +\infty,$$

co też w końcu pociąga $J_N(\pi, X_0) = \infty$ i w konsekwencji (2.10). Zauważmy, że dla strategii $\widehat{\pi}$ mamy równość

$$J_N(\widehat{\pi}, X_0) = \mathbb{E}\, V_0(X_0).$$

Stąd i z pierwszej części twierdzenia wynika optymalność $\widehat{\pi}$. $\Box$

2.3.1 Interpretacja

Jak w przykładzie wprowadzającym, w którym gracz w kasynie chce zmaksymalizować swoją satysfakcję, oznaczmy przez $V_n(x)$ maksymalny średni zysk (minimalny średni koszt) na odcinku czasowym od n do N przy założeniu, że w chwili n proces jest w stanie x. Mamy $V_N(x) = r_N(x)$. Następnie dla $n = N-1, \ldots, 0$,

$$V_n(x) = \sup_{u \in U(x)} \left\{ g_n(x,u) + \mathbb{E}\left(V_{n+1}(X^\pi_{n+1}) | X^\pi_n = x,\ u_n(X^\pi_0, \ldots, X^\pi_n) = u \right) \right\}.$$

Ponieważ, z Lematu 2.1,

$$\mathbb{E}\left(V_{n+1}(X^\pi_{n+1}) | X^\pi_n = x,\ u_n(X^\pi_0, \ldots, X^\pi_n) = u \right) = P^u V_{n+1}(x),$$

więc (V_n) są funkcjami wartości występującymi w twierdzeniu.

2.3.2 Zasada optymalności

Rozważmy przypadek funkcjonału zysku. Załóżmy, że

$$\widehat{\pi}_N = (\widehat{u}_0, \ldots, \widehat{u}_{N-1})$$

jest optymalną strategią dla funkcjonału zysku

$$J_N(\pi, X_0) = \mathbb{E}\left(\sum_{n=0}^{N-1} q_n(X_n, u_n) + r_N(X_N)\right)$$

na odcinku czasowym $[0, N]$. Podzielmy teraz odcinek $[0, N]$ na dwie części $[0, M]$ i $[M, N]$. Na odcinku $[M, N]$ rozważmy funkcjonał

$$J_{M,N}(\pi, X_M) = \mathbb{E}\left(\sum_{n=M}^{N-1} q_n(X_n, u_n) + r_N(X_N)\right).$$

Zauważmy, że strategią optymalną dla $J_{M,N}$ jest strategia

$$\widehat{\pi}_{M,N} = (\widehat{u}_M, \ldots, \widehat{u}_{N-1})$$

będąca częścią strategii $\widehat{\pi}$ optymalnej dla J_N. Ponadto, przy danym X_M optymalny zysk wynosi $\mathbb{E}\, V_{N-M}(X_M)$. Pozostały fragment

$$\widehat{\pi}_{M-1} = (\widehat{u}_0, \ldots, \widehat{u}_{M-1})$$

strategii $\widehat{\pi}$ jest optymalny dla funkcjonału zysku

$$J_M(\pi, X_0) = \mathbb{E}\left(\sum_{n=0}^{M-1} q_n(X_n, u_n) + V_{N-M}(X_M)\right)$$

na odcinku $[0, M]$. Fakt ten nosi nazwę *zasady optymalności* lub *zasady programowania dynamicznego*.

2.4 Przypadek szczególny funkcjonału

Załóżmy teraz, że funkcjonał zysku (lub kosztu) ma postać

$$J_N(\pi, X_0) = \mathbb{E}\left\{\sum_{n=0}^{N-1} \gamma^n q(X_n, u_n) + \gamma^N r(X_N)\right\}, \tag{2.11}$$

gdzie $\gamma \geq 0$, a $q\colon E \times U \to [0, +\infty)$ i $r\colon E \to [0, +\infty)$ są funkcjami mierzalnymi. Poniżej pokażemy, że w tym ważnym przypadku zasada indukcji wstecz Bellmana przyjmuje dużo prostszą postać. W tym celu przyjmijmy oznaczenia

$$\mathcal{A}v(x) = \sup_{u \in U(x)} \left(q(x, u) + \gamma P^u v(x)\right), \tag{2.12}$$

gdy J_N jest funkcjonałem zysku, oraz

$$\mathcal{A}v(x) = \inf_{u \in U(x)} \left(q(x, u) + \gamma P^u v(x)\right), \tag{2.13}$$

gdy J_N jest funkcjonałem kosztu. Niech $\mathcal{A}^j$ oznacza j-tą iteracje operatora $\mathcal{A}$. Przyjmujemy, że $\mathcal{A}^0$ jest operatorem identycznościowym.

Twierdzenie 2.2 *Funkcje wartości* (V_n) *zdefiniowane w Twierdzeniu* 2.1 *dane są formułami*

$$V_n(x) = \gamma^n \mathcal{A}^{N-n} r(x), \qquad n = 0, \dots, N. \tag{2.14}$$

Ponadto, jeżeli $\{\widehat{v}_n\}$ *jest ciągiem funkcji mierzalnych* $E \to U$*, takich że dla dowolnych* n *i* $x \in E$*,* $\widehat{v}_n(x) \in U(x)$ *oraz*

$$\mathcal{A}^{N-n} r(x) = q(x, \widehat{v}_n(x)) + \gamma P^{\widehat{v}_n(x)} \mathcal{A}^{N-(n+1)} r(x), \tag{2.15}$$

to strategia $\widehat{\pi}$ *dana wzorem* (2.8) *jest optymalna.*

Dowód Najpierw wykażemy (2.14) indukcyjnie względem $k = N - n$. Gdy $k = 0$, to $n = N$ i (2.14) zachodzi ze względu na to, że $V_N = \gamma^N r = \gamma^N \mathcal{A}^0 r$. Załóżmy, że (2.14) zachodzi dla pewnego $k = N - n$. Tak więc zakładamy, że $\gamma^{N-k} \mathcal{A}^k r = V_{N-k}$. Wówczas dla funkcjonału zysku (dla kosztu rozumowanie jest podobne) mamy

$$\begin{aligned}
V_{N-k-1}(x) &= \sup_{u \in U(x)} \big(\gamma^{N-k-1} q(x,u) + P^u V_{N-k}(x)\big) \\
&= \sup_{u \in U(x)} \big(\gamma^{N-k-1} q(x,u) + \gamma^{N-k} P^u \mathcal{A}^k r(x)\big) \\
&= \gamma^{N-k-1} \sup_{u \in U(x)} \big(q(x,u) + \gamma P^u \mathcal{A}^k r(x)\big) \\
&= \gamma^{N-k-1} \mathcal{A}^{k+1} r(x).
\end{aligned}$$

Tak więc (2.14) jest udowodnione. Pokażemy teraz, że strategia zadana przez (2.15) jest optymalna. Z zasady indukcji wstecz Bellmana wynika, że w tym celu wystarczy udowodnić równości

$$V_n(x) = \gamma^n q(x, \widehat{v}_n(x)) + P^{\widehat{v}_n(x)} V_{n+1}(x), \qquad x \in E, \ n = 0, \dots, N-1.$$

Wynikają one z (2.14) i (2.15). Istotnie

$$\begin{aligned}
V_n(x) &= \gamma^n \mathcal{A}^{N-n} r(x) \\
&= \gamma^n \Big(q(x, \widehat{v}_n(x)) + \gamma P^{\widehat{v}_n(x)} \mathcal{A}^{N-(n+1)} r(x)\Big) \\
&= \gamma^n q(x, \widehat{v}_n(x)) + P^{\widehat{v}_n(x)} V_{n+1}(x).
\end{aligned}$$

□

2.5 Uogólnienie twierdzenia Bellmana

Gdy mamy do czynienia ze sterowaniem z częściową obserwacją, patrz rozdziały poświęcone problemom sterowania z niepełną obserwacją i filtracji skończonego łańcucha Markowa, potrzebujemy uogólnienia twierdzenia Bellmana na przypadek sterowanego łańcucha Markowa $(X_n^\pi, \mathcal{F}_n)$, patrz Definicja 1.6. Mamy następujący wariant twierdzenia Bellmana.

Twierdzenie 2.3 *Załóżmy, że funkcjonał zysku (kosztu) jest postaci* (2.3) *oraz* V_N, V_{N-1}, ..., V_0 *będą mierzalnymi funkcjami na* E, *takimi że, jeżeli* J_N *jest funkcjonałem zysku to*

$$V_N(x) = r_N(x), \qquad V_n(x) = \sup_{u \in U(x)} \left(q_n(x, u) + P_n^u V_{n+1}(x) \right),$$

a jeżeli J_N *jest funkcjonałem kosztu,*

$$V_N(x) = r_N(x), \qquad V_n(x) = \inf_{u \in U(x)} \left(q_n(x, u) + P_n^u V_{n+1}(x) \right).$$

Wówczas dla dowolnej strategii π, $J_N(X_0^\pi, \pi) \leq \mathbb{E}\, V_0(X_0^\pi)$, *gdy* J_N *jest zyskiem, a* $J_N(X_0^\pi, \pi) \geq \mathbb{E}\, V_0(X_0^\pi)$, *gdy* J_N *jest kosztem. Ponadto, jeżeli istnieją odwzorowania mierzalne* $\widehat{v}_0, \ldots, \widehat{v}_{N-1} \colon E \to U$ *takie, że dla wszystkich* n *i* $x \in E$,

$$V_n(x) = q_n(x, \widehat{v}_n(x)) + P_n^{\widehat{v}_n(x)} V_{n+1}(x),$$

to strategia $\widehat{\pi} = \{\widehat{u}_n\}$, *gdzie*

$$\widehat{u}_n(x_0, x_1, \ldots, x_n) = \widehat{v}_n(x_n)$$

jest optymalna, oraz $J_N(\widehat{\pi}, X_0^\pi) = \mathbb{E}\, V_0(X_0^\pi)$.

Dowód Załóżmy, że J_N jest funkcjonałem zysku. Wówczas dla dowolnych $n = 0, \ldots, N-1$, $x \in E$ oraz $u \in U(x)$ mamy

$$V_n(x) \geq q_n(x, u) + P_n^u V_{n+1}(x).$$

Niech π będzie strategią i niech $X = X^\pi$. Wówczas

$$\begin{aligned} \mathbb{E}\, V_n(X_n) &\geq \mathbb{E}\, q_n(X_n, u_n) + \mathbb{E}\, P_n^{u_n} V_{n+1}(X_n) \\ &\geq \mathbb{E}\, q_n(X_n, u_n) + \mathbb{E}\, V_{n+1}(X_{n+1}). \end{aligned}$$

Teraz możemy postępować jak w dowodzie oryginalnego twierdzenia Bellmana. $\square$

2.6 Przykład

Rozważmy skończone przestrzenie stanów E i sterowań U. Dla $u \in U$ zadana jest macierz przejścia

$$P^u = \left[p_{i,j}^u \right], \qquad i, j \in E.$$

Przypomnijmy, że

$$p_{i,j}^u = \mathbb{P}\left(X_{n+1}^\pi = j | X_n^\pi = i \right) \geq 0,$$

a więc

$$\sum_{j\in E} p^u_{i,j} = 1, \qquad \forall\, i \in E.$$

Załóżmy, że w dowolnym momencie czasowym n, kontroler będąc w stanie $i \in E$ ma prawo wybrać parametr $u \in U$. Wtedy z prawdopodobieństwem $p^u_{i,j}$ w momencie $n+1$ znajdzie się w stanie j. Zauważmy, że dowolna funkcja $h\colon E \mapsto \mathbb{R}$ może być identyfikowana z wektorem $h \in \mathbb{R}^m$ gdzie m jest licznością E. Dokładniej ustawmy wszystkie elementy zbioru E w ciąg $x_1, x_2, \ldots, x_m$. Wtedy

$$h \equiv (h(x_1), \ldots, h(x_m)).$$

Załóżmy teraz, że $E = \{1,2,3\}$ i $U = \{0,1\}$ Niech

$$P^0 = \begin{bmatrix} 0 & \frac{1}{4} & \frac{3}{4} \\ 0 & \frac{1}{4} & \frac{3}{4} \\ 0 & \frac{1}{4} & \frac{3}{4} \end{bmatrix}, \qquad P^1 = \begin{bmatrix} \frac{3}{4} & \frac{1}{4} & 0 \\ \frac{3}{4} & \frac{1}{4} & 0 \\ \frac{3}{4} & \frac{1}{4} & 0 \end{bmatrix}.$$

Z postaci P^0 i P^1 wynika, że jeżeli $u = 0$ to z dużym prawdopodobieństwem skoczymy do stanu 3, a jeśli jesteśmy w stanie 3 to z dużym prawdopodobieństwem zostaniemy w stanie 3. Natomiast jeżeli $u = 1$, to z dużym prawdopodobieństwem skoczymy do stanu 1, a jeśli jesteśmy w stanie 1 to z dużym prawdopodobieństwem zostaniemy w stanie 1. Rozważmy funkcjonał kosztu

$$J_N(x,\pi) = \mathbb{E}\left(\sum_{n=0}^{N-1} u_n + a\chi_{2,3}(X_N)\right).$$

Interpretujemy koszt końcowy $a\chi_{2,3}(X_N)$ jako karę za znalezienie się w niepożądanych położeniach 2 i 3 a koszty bieżące u_n jako opłatę za wybór sterowania perferującego pobyt w pożądanym stanie 1.

Mamy

$$\begin{aligned} V_n(i) &= \min_{u=0,1}\left\{ u + \sum_{j=1,2,3} P^u_{i,j} V_{n+1}(j)\right\} \\ &= \frac{1}{4}\min\left\{V_{n+1}(2) + 3V_{n+1}(3), 4 + 3V_{n+1}(1) + V_{n+1}(2)\right\} \end{aligned}$$

oraz

$$V_N = (0, a, a).$$

Oczywiście dla $n < N$, V_n jest wektorem w $\mathbb{R}^3$ o jednakowych współrzędnych. Mamy

$$V_{N-1}(i) = \min\left\{a, 1 + \frac{a}{4}\right\},$$

a dla $n < N-1$,

$$V_n(i) = \min\left\{V_{n+1}(i), 1 + \frac{V_{n+1}(i)}{4}\right\}.$$

Stąd jeżeli $a \leq 4/3$ to

$$V_{N-1}(i) = V_{N-2}(i) = \ldots = V_0(i) = a,$$

a optymalną strategią jest zawsze wybierać $u = 0$. Natomiast, jeżeli $a > 4/3$, to

$$V_{N-1}(i) = 1 + a/4 > 4/3,$$

więc optymalną stratregią jest zawsze $u = 1$, a

$$V_{N-n}(i) = a_n, \qquad n = 1, 2, \ldots, N, \ i = 1, 2, 3,$$

gdzie

$$a_1 = 1 + \frac{a}{4}, \quad a_{n+1} = 1 + \frac{a_n}{4}.$$

2.7 Problemy inwestora

Przypomnijmy, że w problemie Samuelsona, patrz Przykład 1.1,

$$X_{n+1} = (1 + \xi_{n+1} + b_n(r - \xi_{n+1}))\,(X_n - c_n)$$

jest kapitałem inwestora w chwili $n+1$, przy założeniu, że jego kapitał w chwili n wynosił X_n, w chwili n skonsumował $0 \leq c_n \leq X_n$, $b_n(X_n - c_n)$ ulokował w banku na stałą stopę zwrotu r, a $(1 - b_n)(X_n - c_n)$ ulokował w aktywa ryzykowne, które dają losową stopę zwrotu ξ_{n+1}.

Ustalmy skończony horyzont czasowy N i współczynnik dyskontujący $\gamma > 0$. Celem inwestora jest maksymalizacja oczekiwanej satysfakcji

$$J_N(X_0, \pi) = \mathbb{E}\left\{\sum_{k=0}^{N-1} \gamma^n U(c_n) + \gamma^N \omega U(X_n)\right\},$$

gdzie $U\colon [0, \infty) \to \mathbb{R}$ jest zadaną funkcją satysfakcji a $\omega \geq 0$.

Przypomnijmy, że $U\colon [0, \infty) \to \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$ jest funkcją satysfakcji, gdy jest wklęsła i ścisle rosnąca. Poniżej rozważymy dwa przypadki:

(i) Oryginalny problem Samuelsona; $U(c) = c^\alpha$,
(ii) $U(c) = \log c$.

Zauważmy, że dla $(c, b) \in [0, x] \times [0, 1]$ i $x \geq 0$ mamy

$$P^{(c,b)}h(x) = \mathbb{E}\, h\left(\left[(1 + r)b + (1 - b)(1 + \xi_0)\right](x - c)\right).$$

Przypomnijmy, że operator Bellmana ma postać

$$\mathcal{A}h(x) = \sup_{0 \leq c \leq x,\ b \in [0,1]} \Big(U(c) + \gamma P^{(c,b)}h(x)\Big).$$

2.7.1 Klasyczny problem Samuelsona; $U(c) = c^\alpha$

Niech

$$\widehat{\lambda} = \sup_{b\in[0,1]} \mathbb{E}\left((1+r)b + (1-b)(1+\xi_0)\right)^\alpha . \tag{2.16}$$

Załóżmy też, że supremum w (2.17) jest osiągane w punkcie $\widehat{b} \in [0,1]$. Dla $z \geq 0$ połóżmy

$$\widehat{f}(z) = \left(1 + (z\gamma\widehat{\lambda})^{1/(1-\alpha)}\right)^{1-\alpha}, \quad \widehat{g}(z) = \frac{1}{1 + (z\gamma\widehat{\lambda})^{1/(1-\alpha)}}.$$

W poniższym twierdzeniu będziemy zakładali, że $\alpha \in (0,1)$. Przypadki $\alpha = 0, 1$ są dużo łatwiejsze i pozostawiamy je czytelnikowi.

Twierdzenie 2.4 *Dla problemu inwestora funkcje wartości $\{V_n\}$ zdefiniowane w Twierdzeniu 2.1 dane są wzorami*

$$V_n(x) = \gamma^n \widehat{f}^{N-n}(\omega) x^\alpha, \qquad n = 0, \ldots, N. \tag{2.17}$$

Ponadto optymalne sterowania $\widehat{u}_n = (\widehat{c}_n, \widehat{b}_n)$ dane są formułami

$$\widehat{c}_n(x_0, \ldots x_n) = \widehat{g}\left(\widehat{f}^{N-(n+1)}(\omega)\right) x_n, \qquad n = 0, \ldots, N-1,$$
$$\widehat{b}_n(x_0, \ldots, x_n) = \widehat{b}, \qquad n = 0, \ldots, N-1.$$

Dowód Pokażemy, że dla operatora Belmana $\mathcal{A}$ oraz funkcji $r_z(x) := zx^\alpha$, $z, x \geq 0$, mamy $\mathcal{A}r_z(x) = \widehat{f}(z)x^\alpha$. Istotnie, z definicji operatora $\mathcal{A}$,

$$\begin{aligned}
&\mathcal{A}r_z(x) \\
&= \sup_{0\leq c\leq x,\ b\in[0,1]} \left(c^\alpha + \gamma P^{(c,b)} r_z(x)\right) \\
&= \sup_{0\leq c\leq x,\ b\in[0,1]} \left(c^\alpha + \gamma (x-c)^\alpha z \mathbb{E}\left((1+r)b + (1-b)(1+\xi_0)\right)^\alpha\right) \\
&= \sup_{0\leq c\leq x} \left(c^\alpha + \gamma(x-c)^\alpha z\widehat{\lambda}\right).
\end{aligned}$$

Skorzystamy teraz z następującego elementarnego faktu. Mianowicie dla dowolnych $z \geq 0$ i $x \geq 0$ zachodzi

$$\sup_{0\leq c\leq x} \left[c^\alpha + \gamma(x-c)^\alpha z\widehat{\lambda}\right] = x^\alpha \widehat{f}(z),$$

przy czym supremum jest osiągane w punkcie $\widehat{c} = \widehat{g}(z)$. Tak więc

$$\mathcal{A}r_z(x) = \widehat{f}(z)x^\alpha.$$

Ponadto

$$\mathcal{A}r_z(x) = (\widehat{g}(z))^\alpha + \gamma P^{\left(\widehat{g}(z),\widehat{b}\right)} r_z(x). \tag{2.18}$$

Stąd
$$\mathcal{A}^n r_z(x) = \widehat{f}^n(z)x^\alpha,$$
a więc
$$\mathcal{A}^n r(x) = \widehat{f}^n(\omega)x^\alpha$$
oraz, z Twierdzenia 2.2, otrzymujemy (2.17). Należy wiec wykazać, że ciąg $\{\widehat{u}_n\}$ spełnia (2.15). Wynika to z (2.18) oraz z następującego rachunku. Dla

$$\widehat{v}(x) = \widehat{u}(x_0, \ldots, x_{n-1}, x)$$

zachodzi

$$\begin{aligned} q(x, \widehat{v}_n(x)) &+ \gamma P^{\widehat{v}_n(x)} \mathcal{A}^{N-(n+1)} r(x) \\ &= q(x, \widehat{v}_n(x)) + \gamma P^{\widehat{v}_n(x)} \widehat{f}^{N-(n+1)}(\omega) r_1(x) \\ &= \mathcal{A}\, r_{\widehat{f}^{N-(n+1)}(\omega)}(x) \\ &= \mathcal{A}^{N-n}\, r(x). \end{aligned}$$

□

2.7.2 Problem inwestora z logarytmiczną funkcją satysfakcji; $U(c) = \log c$

Rozważmy przypadek gdy $U(c) = \log c$. Niech

$$r_{z,v}(x) = z + v \log x.$$

Mamy

$$\begin{aligned} \mathcal{A} r_{z,v}(x) &= \sup_{c\in[0,x],\, b\in[0,1]} \left\{ \log c + \gamma P^{(c,b)} r_{z,v}(x) \right\} \\ &= \sup_{c\in[0,x],\, b\in[0,1]} \{\log c + \gamma z + \gamma v \mathbb{E} \log\left[(1 + \xi_1 + b(r - \xi_1))\,(x - c)\right]\} \\ &= \gamma(z + v\widehat{\lambda}) + \sup_{c\in[0,x]} \{\log c + \gamma v \log(x - c)\}, \end{aligned}$$

gdzie
$$\widehat{\lambda} := \sup_{b\in[0,1]} \mathbb{E} \log(1 + \xi_1 + b(r - \xi_1)).$$

Załóżmy, że supremum jest osiągane w $\widehat{b}$.

Tak więc

$$\mathcal{A} r_{z,v}(x) = \gamma(z + v\widehat{\lambda}) + \sup_{c\in[0,x]} \{\log c + \gamma v \log(x - c)\}$$

Oczywiście supremum jest osiągane dla

$$\widehat{c} := \frac{x}{1+\gamma v}.$$

Wynosi ono

$$\log \frac{x}{1+\gamma v} + \gamma v \log \frac{\gamma v x}{1+\gamma v} = (1+\gamma v)\log x + \gamma v \log \gamma v - (1+\gamma v)\log(1+\gamma v).$$

Stąd

$$\mathcal{A} r_{z,v}(x) = \gamma(z + v\widehat{\lambda}) + \gamma v \log \gamma v - (1+\gamma v)\log(1+\gamma v) + (1+\gamma v)\log x.$$

Niech

$$f(z,v) = \gamma(z + v\widehat{\lambda}) + \gamma v \log \gamma v - (1+\gamma v)\log(1+\gamma v),$$
$$g(v) = \frac{1}{1+\gamma v}.$$

Mamy więc

$$\mathcal{A} r_{z,v} = r_{f(z,v),g(v)}.$$

Ponadto

$$(g(v)x, \widehat{b}) = \operatorname{Arg} \sup_{c\in[0,x],\, b\in[0,1]} \left\{ \log c + \gamma P^{(c,b)} r_{z,v}(x) \right\}.$$

Mamy więc następujący wynik.

Twierdzenie 2.5 *Dla problemu inwestora z logarytmiczną funkcją satysfakcji funkcje wartości* $\{V_n\}$ *zdefiniowane w Twierdzeniu* 2.1 *dane są wzorami*

$$V_n(x) = \gamma^n \mathcal{A}^{N-n} r_{0,\omega}(x) = \gamma^n \left(z_{N-n} + \omega_{N-n} \log x \right), \qquad n = 0, \ldots, N,$$

gdzie $(z_0, \omega_0) = (0, \omega)$ *a*

$$(z_{n+1}, \omega_{n+1}) = \left(f(z_n, \omega_n), h(\omega_n) \right), \qquad n = 0, 1, \ldots.$$

Ponadto optymalne sterowania $\widehat{u}_n = (\widehat{c}_n, \widehat{b}_n)$ *dane są formułami*

$$\widehat{c}_n(x_0, \ldots x_n) = \widehat{g}\left(\omega_{N-(n+1)}\right) x_n, \qquad n = 0, \ldots, N-1,$$
$$\widehat{b}_n(x_0, \ldots, x_n) = \widehat{b}, \qquad n = 0, \ldots, N-1.$$

2.8 Problem maksymalizacji końcowego kapitału

Załóżmy, że inwestor nie konsumuje żadnej części swojego kapitału X_n w chwili n, ale podobnie jak w w problemie Samuelsona, patrz Przykład 1.1, może inwestować $b_n X_n$ w banku na stałą stopę zwrotu r i $(1-b_n)X_n$ w aktywa ryzykowne z losową stopą zwrotu ξ_{n+1}. Zakładamy, że ξ_n są niezależnymi zmiennymi losowymi o tym samym rozkładzie oraz, że $\mathbb{P}(\xi_n \geq -1) = 1$. Tak więc dynamika kapitału inwestora dana jest wzorem

$$X_{n+1} = (1 + \xi_{n+1} + b_n(r - \xi_{n+1}))\, X_n.$$

Ustalmy skończony horyzont czasowy N. Celem inwestora jest maksymalizacja oczekiwanej satysfakcji z kapitału końcowego X_N;

$$J_N(X_0, \pi) = \mathbb{E}\, U(X_n),$$

gdzie $U \colon [0, \infty) \to \mathbb{R}$ jest zadaną funkcją satysfakcji.

Poniżej rozważymy dwa przypadki:

(i) $U(x) = x^\alpha$,
(ii) $U(x) = \log x$.

2.8.1 Przypadek $U(x) = x^\alpha$

Niech $r_z(x) = zx^\alpha$. Mamy, patrz Rozdział 2.7.1,

$$\begin{aligned}
\mathcal{A}r_x(x) &= \sup_{b\in[0,1]} P^{(b)} r_z(x) \\
&= \sup_{b\in[0,1]} z\mathbb{E}\left[(1 + \xi_1 + b(r - \xi_1))\, x\right]^\alpha \\
&= z\widehat{\lambda} x^\alpha = r_{\widehat{\lambda} z}(x),
\end{aligned}$$

gdzie

$$\widehat{\lambda} := \sup_{b\in[0,1]} \mathbb{E}\left(1 + \xi_1 + b(r - \xi_1)\right)^\alpha.$$

Niech

$$\widehat{b} := \operatorname{Arg} \sup_{b\in[0,1]} \mathbb{E}\left(1 + \xi_1 + b(r - \xi_1)\right)^\alpha.$$

Mamy więc następujący wynik:

Twierdzenie 2.6 *Optymalną strategią jest* $\widehat{\pi} = \left(\widehat{b}_0, \ldots, \widehat{b}_{N-1}\right)$, *gdzie*

$$\widehat{b}_n(x_0, \ldots, x_n) = \widehat{b}.$$

Ponadto

$$J_N(X_0, \widehat{\pi}) = \left(\widehat{\lambda}\right)^N \mathbb{E}\left(X_0\right)^\alpha.$$

2.8.2 Przypadek $U(x) = \log x$

W tym przypadku dla funkcji

$$r_z(x) = z + \log x.$$

mamy

$$
\begin{aligned}
\mathcal{A}r_z(x) &= \sup_{b\in[0,1]} P^{(b)} r_z(x) \\
&= z + \sup_{b\in[0,1]} \mathbb{E}\log\left[(1+\xi_1+b(r-\xi_1))\,x\right] \\
&= z + \widehat{\lambda} + \log x = r_{z+\widehat{\lambda}}(x),
\end{aligned}
$$

gdzie

$$
\widehat{\lambda} := \sup_{b\in[0,1]} \mathbb{E}\log\left(1+\xi_1+b(r-\xi_1)\right).
$$

Niech

$$
\widehat{b} = \mathrm{Arg}\sup_{b\in[0,1]} \mathbb{E}\log\left(1+\xi_1+b(r-\xi_1)\right).
$$

Mamy

Twierdzenie 2.7 *Optymalną strategią jest* $\widehat{\pi} = \left(\widehat{b}_0, \ldots, \widehat{b}_{N-1}\right)$*, gdzie*

$$
\widehat{b}_n(x_0, \ldots, x_n) = \widehat{b}.
$$

Ponadto

$$
J_N(X_0, \widehat{\pi}) = N\widehat{\lambda} + \mathbb{E}\,\log X_0.
$$

2.8.3 Problem inwestora z proporcjonalmymi i stałymi kosztami za transakcje

Załóżmy, że inwestor kupując akcje warte y musi zapłacić proporcjonalne koszty transakcyjne λy i stałe koszty a, a więc wydać $(1+\lambda)y + a$ z kapitału ulokowanego w banku. Podobnie spieniężając kapitał y ulokowany w akcjach otrzymuje $(1-\mu)y - b$ na rachunku bankowym płacąc μy proporcjonalnych kosztów za transakcję i b kosztów stałych.

Przestrzenią stanów w problemie z kosztami za transakcje jest $E = \mathbb{R}^2$. Potrzebujemy dwóch parametrów do opisu. Mianowicie, niech X_n i Y_n oznaczają kapitały zainwestowane odpowiednio w banku i na rachunku akcyjnym. Dopuszczamy, ujemne wartości X_n lub Y_n. Wymagamy jednak by inwestor był wypłacalny to znaczy by $(X_n, Y_n) \in \mathcal{C}$, gdzie obszar wypłacalności $\mathcal{C}$ dany jest następująco

$$
\begin{aligned}
\mathcal{C} = (0,+\infty)^2 &\cup \left\{(x,y) \in \mathbb{R}\times(0,+\infty) \colon x + (1-\mu)y - b \geq 0\right\} \\
&\cup \left\{(x,y) \in (0,+\infty)\times\mathbb{R} \colon x + (1+\lambda)y + a \geq 0\right\}.
\end{aligned}
$$

Dynamika (X_n, Y_n) jest następująca

$$
\begin{aligned}
X_{n+1} &= (1+r)\left[X_n + (1-\mu)m_n - b\chi_{(0,+\infty)}(m_n) - (1+\lambda)l_n \right.\\
&\qquad \left. - a\chi_{(0,+\infty)}(l_n) - c_n\right], \\
Y_{n+1} &= (1+\xi_{n+1})(Y_n - m_n + l_n),
\end{aligned}
$$

gdzie $r \in \mathbb{R}$ jest stałą stopą oferowaną przez bank (na depozyty i pożyczki), a (ξ_n) jest ciągiem niezależnych zmiennych losowych o tym samym rozkładzie spełniającym $\mathbb{P}(\xi_n \geq -1) = 1$. Jak w oryginalnym problemie Samuelsona konsumpcję onaczamy przez c_n. W modelu z opłatami za transakcje alokacja kapitału jest bardziej złożona niż w modelu bez kosztów. Mianowicie przez m_n oznaczamy wartość akcji, którą przeznaczamy do sprzedaży w chwili n, a przez l_n oznaczamy wartość akcji kupionych w chwili n. Zakładamy, że $l_n, m_n \geq 0$, przy czym ich wartości ściśle dodatnie powodują występowanie odpowiednich kosztów stałych.

Niestety problem z kosztami za transakcje nie ma znanego w postaci analitycznej rozwiązania! Więcej o tym problemie można znaleźć w pracy: Roman V. Bobryk i Łukasz Stettner, Discrete Time Portfolio Selection with Proportional Transaction Costs, Probab. Math. Statist. 19 (1999), 235–248.

2.9 Przykłady

2.9.1 Problem błądzenia z dryfem

Niech $E = \{\ldots -3, -2, -1, 0\}$, $U = \{0, 1\}$. Mamy sterowany łańcuch Markowa X na E z prawdopodobieństwami przejścia

$$p^0_{i,i+1} = 0.5 = p^0_{i,i-1}, \qquad p^1_{i,i+1} = 0.75, \quad p^1_{i,i-1} = 0.25, \qquad i < 0,$$

$p^i_{0,0} = 1$, $i \in U$. Naszym celem jest dosterowania go, na skończonym przedziale czasowym $[0, N]$, jak najbliżej punktu pochłaniającego 0. Wyliczymy funkcję wartości i optymalne sterowanie dla horyzontu czasowego $N = 2$ i funkcjonału kosztu kosztu

$$J\left((u_n), X_0\right) = \mathbb{E}\left(\sum_{n=0}^{N-1} u_n + X_N^2\right).$$

Mamy

$$q(x,u) = u, \qquad r(x) = x^2.$$

Ponadto

$$P^0 v(x) = \begin{cases} \frac{1}{2}\left(v(x-1) + v(x+1)\right) & \text{gdy } x < 0, \\ v(0) & \text{gdy } x = 0 \end{cases}$$

oraz

$$P^1 v(x) = \begin{cases} \frac{1}{4} v(x-1) + \frac{3}{4} v(x+1) & \text{gdy } x < 0, \\ v(0) & \text{gdy } x = 0. \end{cases}$$

Stąd, dla $x < 0$,

$$\mathcal{A}v(x) = \min\left\{\frac{1}{2}\left(v(x-1) + v(x+1)\right), 1 + \frac{1}{4} v(x-1) + \frac{3}{4} v(x+1)\right\}.$$

Czyli, dla $x < 0$,

$$\mathcal{A}r(x) = \min\{x^2+1, x^2+x+2\} = x^2+x+2.$$

Oczywiście $\mathcal{A}v(0) = v(0)$. Stąd

$$\mathcal{A}r(x) = \begin{cases} x^2+x+2 & \text{dla } x<0, \\ 0 & \text{dla } x=0 \end{cases}$$

oraz sterowanie u_{N-1} dane jest wzorem

$$u_{N-1}(x) = \begin{cases} 1 & \text{dla } x<0, \\ 0 & \text{dla } x=0. \end{cases}$$

Zauważmy, że ponieważ $P^0(-i) = P^1r(-1)$, więc możemy również przyjąć $u_{N-1}(-1) = 0$. Oczywiście mamy $\mathcal{A}r(0) = 0 = \mathcal{A}^jr(0)$. Następnie, dla $x < -1$ mamy

$$\mathcal{A}^2r(x) = \min\left\{x^2+1+2, x^2+2x+4-\frac{1}{4}+\frac{3}{4}\right\} = x^2+2x+4+\frac{1}{2}.$$

Natomiast dla $x=-1$,

$$\begin{aligned}\mathcal{A}^2r(x) &= \min\left\{\frac{1}{2}(x-1)^2+\frac{1}{2}(x-1)+2, \frac{1}{4}(x-1)^2+\frac{1}{4}(x-1)+3\right\} \\ &= \frac{1}{2}(x-1)^2+\frac{1}{2}(x-1)+2.\end{aligned}$$

Stąd

$$u_{N-2}(x) = \begin{cases} 1 & \text{dla } x<-1, \\ 0 & \text{dla } x=-1,0. \end{cases}$$

Podsumowując, w chwili $n=0$ powinniśmy wybrać sterowanie $u=1$ gdy jesteśmy w położeniu $x \le -2$ a 0 gdy w położeniu $x=-1$ lub $x=0$. W chwili 1 powinniśmy wybrać sterowanie $u=1$ gdy jesteśmy w położeniu $x \le -1$ i 0 gdy jesteśmy w położeniu 0. Funkcją wartości jest

$$\mathcal{A}^2r(x) = \begin{cases} x^2+2x+4+\frac{1}{2} & \text{dla } x \le -2, \\ 6 & \text{dla } x=-1, \\ 0 & \text{dla } x=0. \end{cases}$$

2.9.2 Problem śledzenia

Niech $\{Z_n\}$ będzie sterowanym łancuchem Markowa na $\mathbb{R}$ zadanym rekurencyjnie

$$Z_{n+1} = \alpha Z_n + u_n + \xi_{n+1}.$$

Zakładamy, że $\{\xi_n\}$ jest ciągiem niezależnych zmiennych losowych o jednakowym rozkładzie na $\mathbb{R}$. Celem jest jak najbliższe dosterowanie Z do procesu

$$C_{n+1} = \beta C_n + \eta_{n+1}.$$

Tutaj $\{\eta_n\}$ jest ciągiem iid o wartościach w $\mathbb{R}$ niezależnym od $\{\xi_n\}$. Sterowanie u_n może zależeć od obserwowalnych wielkości Z_n i C_n.

Zakładamy, że koszty związane ze sterowaniem wynoszą

$$J((u_n), Z_0, C_0) = \mathbb{E}\left\{ \sum_{n=0}^{N-1} q(u_n) + h(Z_N - C_N) \right\}.$$

Sformułujemy równania Bellmana, a następnie znajdziemy optymalną strategię dla $q(u) = u^2 = h(x)$, $\alpha = 1 = -\beta$, przy założeniu, że ξ_n i η_n mają średnie 0.

Przestrzenią stanów jest $E = \mathbb{R} \times \mathbb{R}$, a przestrzenią sterowań jest $U = \mathbb{R}$. Operator przejścia dany jest wzorem

$$\begin{aligned} P^u v(z,c) &= \mathbb{E} v(\alpha z + u + \xi_0, \beta c + \eta_0) \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} v(\alpha z + u + x, \beta c + y) \psi_\xi(\mathrm{d}x) \psi_\eta(\mathrm{d}y), \end{aligned}$$

gdzie ψ_ξ i ψ_η są rozkładami ξ_0 i η_0. Operator Bellmana ma postać

$$\mathcal{A} v(z,c) = \inf_{u \in \mathbb{R}} \left\{ q(u) + P^u v(z,c) \right\}.$$

Rozważmy przypadek gdy $q(u) = u^2$ i $r(z,c) = h(z-c) = (z-c)^2$ i $\alpha = 1 = -\beta$. Wtedy, ze faktu, że ξ_n i η_n mają średnie zero i są niezależne mamy

$$\begin{aligned} \mathcal{A} r(z,c) &= \inf_{u \in \mathbb{R}} \left\{ u^2 + \mathbb{E}\left((z + u + c) + (\xi_0 - \eta_0) \right)^2 \right\} \\ &= \mathbb{E}\left(\xi_0^2 + \eta_0^2\right) + \inf_{u \in \mathbb{R}} \left\{ u^2 + (z + c + u)^2 \right\} \\ &= \mathbb{E}\left(\xi_0^2 + \eta_0^2\right) + \frac{(z+c)^2}{2}, \end{aligned}$$

przy czym minimum jest osiągane dla

$$u_{N-1} = -\frac{z+c}{2}.$$

Następnie

$$\begin{aligned} \mathcal{A}^2 r(z,c) &= \frac{3}{2} \mathbb{E}\left(\xi_0^2 + \eta_0^2\right) + \inf_{u \in \mathbb{R}} \left\{ u^2 + \frac{(z + u - c)^2}{2} \right\} \\ &= \frac{3}{2} \mathbb{E}\left(\xi_0^2 + \eta_0^2\right) + \frac{(z-c)^2}{3}, \end{aligned}$$

przy czym minimum jest osiągane dla

$$u_{N-2} = -\frac{z-c}{3}.$$

Stąd

$$\mathcal{A}^n r(z,c) = a_n \, \mathbb{E}\left(\xi_0^2 + \eta_0^2\right) + \frac{\left(z + (-1)^{n+1} c\right)^2}{n+1}$$

gdzie, $a_0 = 0$ oraz

$$a_{n+1} = a_n + \frac{1}{n+1},$$

a optymalną strategią na przedziale skończonym $[0, N]$ jest

$$u_n = -\frac{z + (-1)^{N-n+1} c}{N-n+1}, \qquad n = 0, \ldots, N-1,$$

2.10 Sprowadzenie funkcjonału do postaci standardowej

Załóżmy, że na przestrzeni mierzalnej $(E, \mathcal{E})$ zadany jest rekurencyjnie sterowany łańcuch Markowa

$$X_{n+1}^\pi = F(X_n^\pi, u_n(X_0^\pi, \ldots, X_n^\pi), \xi_{n+1}).$$

Niech $M, K \in \mathbb{N}$, $K \leq M$, $A \in \mathcal{E}$. Niech $k_0, \ldots, k_M$, f będą funkcjami mierzalnymi na E oraz niech $a \in \mathbb{R}$. Rozważmy następujące funkcjonały

(a) $\mathbb{E} \min_{0 \leq j \leq M} f(X_j^\pi)$.
(b) $\mathbb{E} \max_{0 \leq j \leq M} f(X_j^\pi)$.
(c) $\mathbb{P}\left(\exists\, j \in \{0, 1, \ldots, M\} \colon X_j^\pi \in A\right)$.
(d) $\mathbb{P}\left(\forall\, j \in \{0, 1, \ldots, M\} \colon X_j^\pi \in A\right)$.
(e)

$$\mathbb{P}\left(\sum_{n=0}^{M-1} k_n(X_n^\pi) \geq a\right) + \mathbb{E}\, k_M(X_M^\pi).$$

(f)

$$\mathbb{P}\left(\max_{j \leq M} f(X_j^\pi) \geq a\right).$$

(g)

$$\mathbb{P}\left(X_j^\pi \in A \text{ dla dokładnie } K\text{-jotów} \leq M\right).$$

(h) Funkcjonał średniego czasu dojścia do zbioru A, to znaczy

$$\mathbb{E}\, \tau_A \qquad \text{i} \qquad \mathbb{E}\, \tau_A \wedge M,$$

gdzie

$$\tau_A = \inf\{n \colon X_n^\pi \in A\}.$$

Pokażemy, że zmieniając przestrzeń stanów można sprowadzić każdy z powyższych funkcjonałów do postaci standartowej

$$J(\overline{\pi}, \overline{X}_0) = \mathbb{E}\left\{\sum_{n=0}^{N-1} q_n(\overline{X}_n, \overline{u}_n) + r_N(\overline{X}_N)\right\},$$

gdzie $\overline{X} := \overline{X}^{\overline{\pi}}$ jest odpowiednio skonstruowanym, zadanym rekurencyjnie sterowanym procesem Markowa oraz $N < +\infty$. W dalszym ciągu będziemy omijać wskaźniki π i $\overline{\pi}$.
Funkcjonał (a). Niech

$$Y_n := \min\{f(X_j) \colon j \le n\}, \qquad n = 0, 1, \dots.$$

Wówczas $\overline{X}_n := (X_n, Y_n)$ jest sterowanym procesem Markowa zadanym rekurencyjnie przez $Y_0 = f(X_0)$ i

$$Y_{n+1} = \min\{Y_n, f(X_{n+1})\} = \min\{Y_n, f(F(X_n, u_n, \xi_{n+1}))\}.$$

Oczywiście

$$\mathbb{E} \min_{0\le j\le M} f(X_j) = \mathbb{E}\, Y_M.$$

Stąd

$$J(\overline{\pi}, \overline{X}_0) = \mathbb{E}\, r(\overline{X}_M),$$

gdzie $r_M(x, y) = y$.
Funkcjonał (b). Stosujemy rozumowanie jak dla funkcjonału (a), kładąc

$$Y_n := \max\{f(X_j) :\ j \le n\}, \qquad n \ge 0.$$

Funkcjonał (c). Ponieważ

$$\mathbb{P}\,(\exists\, j \in \{0, 1, \dots, M\} \colon X_j \in A) = \mathbb{E} \max_{0\le j\le M} \chi_A(X_j),$$

funkcjonał (c) jest szczególnym przypadkiem funkcjonału (b).
Funkcjonał (d). Ponieważ

$$\mathbb{P}\,(\forall\, j \in \{0, 1, \dots, M\} \colon X_j \in A) = \mathbb{E} \min_{0\le j\le M} \chi_A(X_j),$$

funkcjonał jest szczególnym przypadkiem funkcjonału (a).
Funkcjonał (e). Zdefiniujmy

$$Y_n := \sum_{j=0}^{n} k_j(X_j), \qquad n \ge 0.$$

Ponieważ

$$Y_{n+1} = Y_n + k_{n+1}(X_{n+1}) = Y_n + k_{n+1}(F(X_n, u_n, \xi_{n+1})),$$

$\overline{X} := (X, Y)$ jest procesem Markowa. Ponadto

$$\begin{aligned}\mathbb{P}\left(\sum_{j=0}^{M-1} k_j(X_j) \geq a\right) &= \mathbb{P}\left(Y_{M-1} \geq a\right) \\ &= \mathbb{E}\,\chi_{[a,+\infty)}(Y_{M-1}).\end{aligned}$$

Funkcjonał (f). Ponieważ

$$\begin{aligned}\mathbb{P}\left(\max_{0\leq j\leq M} f(X_j) \geq a\right) &= \mathbb{P}\left(\exists j \in \{0, 1, \ldots, M\} \colon X_j \in f^{-1}([a, \infty))\right) \\ &= \mathbb{E}\max_{0\leq j\leq M} \chi_{f^{-1}([a,+\infty))}(X_j),\end{aligned}$$

problem sprowadza się do problemów (c) i (b).
Funkcjonał (g). Ponieważ

$$\mathbb{P}\left(X_j \in A \text{ dla dokładnie } K\text{-jotów} \leq M\right) = \mathbb{P}\left(\sum_{j=0}^{M} \chi_A(X_j) = K\right)$$

problem sprowadza się do problemu (e).
Funkcjonał (h). Niech $Y_0 := \chi_A(X_0)$ oraz

$$Y_{n+1} := \max\{\chi_A(X_{n+1}), Y_n\} = \max\{\chi_A(F(X_n, u_n, \xi_{n+1}), Y_n\}.$$

Oczywiście $\overline{X}_n := (X_n, Y_n)$ jest procesem Markowa. Ponadto

$$\mathbb{E}\,\tau_A = \mathbb{E}\sum_{n=0}^{+\infty}(1 - Y_n)$$

oraz

$$\mathbb{E}\,\tau_A \wedge M = \mathbb{E}\sum_{n=0}^{M}(1 - Y_n).$$

2.11 Funkcjonał Markowitza

Harry M. Markowitz, amerykański ekonomista, laureat Nagrody Banku Szwecji im. Alfreda Nobla w dziedzinie ekonomii w 1990 roku, w artykule Portfolio Selection, J. Finance 7 (1952), 77–91, zaproponował następujący problem: inwestor konstruując swój portfel chce zmaksyalizować kapitał końcowy minimalizując ryzyko. W tym celu poszukuje strategii inwestycyjnej $\widehat{\pi}$, dla której oczekiwany kapitał $X_N^{\widehat{\pi}}$ w chwili N zakończenia inwestycji jest duży a wariancja

$$\mathbb{D}^2 X_N^{\widehat{\pi}} := \mathbb{E}\left(X_N^{\widehat{\pi}}\right)^2 - \left(\mathbb{E}\, X_N^{\widehat{\pi}}\right)^2$$

mała. W rezultacie funkcjonał zysku Markowitza ma postać

$$J_N(x,\pi) = \mathbb{E}\, X_N^\pi - \theta \mathbb{D}^2 X_N^\pi,$$

gdzie x jest kapitałem początkowym a $\theta > 0$ jest współczynnikiem związanym z awersją do ryzyka inwestora. Oczywiście funkcjonał zysku nie jest postaci wymaganej w twierdzeniu Bellmana. Pokażemy jak do rozwiązania problemu maksymalizacji funkcjonału Markowitza zastosować zasadę indukcji wstecz Bellmana.

Niech Π oznacza ogół strategii dopuszczalnych inwestora. Mamy

$$\begin{aligned}
\sup_{\pi\in\Pi} J_N(x,\pi) &= \sup_{A\in\mathbb{R}} \sup_{\pi\in\Pi} \{J_N(x,\pi) \colon X_N^\pi = A\} \\
&= \sup_{A\in\mathbb{R}} \sup_{\pi\in\Pi} \left\{A + \theta A^2 - \theta \mathbb{E}\,(X_N^\pi)^2 : X_N^\pi = A\right\} \\
&= \sup_{A\in\mathbb{R}} \left[A + \theta A^2 - \theta \inf_{\pi\in\Pi} \left\{\mathbb{E}\,(X_N^\pi)^2 : \mathbb{E}\, X_N^\pi = A\right\}\right].
\end{aligned}$$

By policzyć

$$\inf_{\pi\in\Pi} \left\{\mathbb{E}\,(X_N^\pi)^2 : \mathbb{E}\, X_N^\pi = A\right\}$$

zastosujemy metodę Lagrange'a liczenia ekstremów warunkowych. Mianowicie dla $\lambda \in \mathbb{R}$ rozważmy problem

$$\begin{aligned}
\inf_{\pi\in\Pi} \left\{\mathbb{E}\,(X_N^\pi)^2 + \lambda\,(\mathbb{E}\, X_N^\pi - A)\right\} &= \inf_{\pi\in\Pi} \left\{\mathbb{E}\,(X_N^\pi + \lambda/2)^2 - \lambda A - \lambda^2/4\right\} \\
&= -\lambda A - \lambda^2/4 + \inf_{\pi\in\Pi} J_N^\lambda(x,\pi) \\
&= A^2 - (A+\lambda/2)^2 + \inf_{\pi\in\Pi} J_N^\lambda(x,\pi)
\end{aligned}$$

gdzie

$$J_N^\lambda(x,\pi) := \mathbb{E}\,(X_N^\pi + \lambda/2)^2\,. \tag{2.19}$$

Załóżmy, że dla ustalonego λ potrafimy znaleźć strategie optymalną $\widehat{\pi}^\lambda$, która minimalizuje $J_N^\lambda(x,\pi)$ po wszystkich dopuszczalnych strategiach $\pi \in \Pi$. Jeżeli teraz znajdziemy takie λ_A, że $\mathbb{E}\, X^{\widehat{\pi}^{\lambda_A}} = A$, to mamy

$$\begin{aligned}
&\inf_{\pi\in\Pi} \left\{\mathbb{E}\,(X_N^\pi)^2 : \mathbb{E}\, X_N^\pi = A\right\} \\
&\quad = \inf_{\pi\in\Pi} \left\{\mathbb{E}\,(X_N^\pi)^2 + \lambda_A\,(\mathbb{E}\, X_N^\pi - A) : \mathbb{E}\, X_N^\pi = A\right\} \\
&\quad = A^2 - (A+\lambda/2)^2 + \inf_{\pi\in\Pi\colon \mathbb{E}\, X_N^\pi = A} J_N^{\lambda_A}(x,\pi).
\end{aligned}$$

Ponieważ dla naszego λ_A mamy

$$\inf_{\pi\in\Pi\colon \mathbb{E}\, X_N^\pi = A} J_N^{\lambda_A}(x,\pi) = \inf_{\pi\in\Pi} J_N^{\lambda_A}(x,\pi),$$

więc

$$\inf_{\pi\in\Pi}\left\{\mathbb{E}\,(X_N^\pi)^2 : \mathbb{E}\,X_N^\pi = A\right\} = A^2 - (A+\lambda/2)^2 + \inf_{\pi\in\Pi} J_N^{\lambda_A}(x,\pi).$$

W rezultacie problem z funkcjonałem zysku Markowitza rozwiązujemy następująco:

(i) Dla ustalonego $\lambda \in \mathbb{R}$ rozwiązujemy klasyczny problem sterowania z funkcjonałem kary $J_N^\lambda(x,\pi)$ danym przez (2.19). Niech

$$f(x,\lambda) := \inf_{\pi\in\Pi} J_N^\lambda(x,\pi)$$

oraz niech $\widehat{\pi}^\lambda$ będzie strategią optymalną.

(ii) Szukamy takiego λ_A, że

$$\mathbb{E}\,X_N^{\widehat{\pi}^{\lambda_A}} = A.$$

(iii) Niech

$$G(A,x) := A + \theta\,(A+\lambda_A/2)^2 - \theta f(x,\lambda_A)$$

oraz niech $\widehat{A}$ będzie argumentem dla którego funkcja $A \mapsto G(A,x)$ osiąga maksimum.

(iv) Optymalną strategią jest $\widehat{\pi} := \widehat{\pi}^{\lambda_{\widehat{A}}}$ oraz

$$\sup_{\pi\in\Pi}\left(\mathbb{E}\,X_N^\pi - \theta\,\mathbb{D}\,X_N^\pi\right) = \mathbb{E}\,X_N^{\widehat{\pi}} - \theta\,\mathbb{D}\,X_N^{\widehat{\pi}} = G(\widehat{A},x).$$

2.11.1 Problem gracza w kasynie z funkcjonałem Markowitza

Rozważmy gracza, który ma nieograniczony dostęp do kredytu i który wygrywa z prawdopodobieństwem p a przegrywa z prawdopodobieństwem $q = 1-p$. Dokładniej niech X_n oznacza kapitał gracza w chwili n. Niech b_n oznacza kwotę postawioną w chwili n. Zakładamy, że b_n może przyjmować wartości ujemne. Wówczas

$$\mathbb{P}\,(X_{n+1} = X_n + b_n) = p, \qquad \mathbb{P}\,(X_{n+1} = X_n - b_n) = 1-p.$$

Zauważmy, że w przypadku symetrycznym $p = q = 1/2$, $\mathbb{E}\,X_N^\pi = X_0^\pi = x$ dla dowolnej strategii π. Stąd w tym przypadku rozwiązanie problemu maksymalizacji funkcjonału Markowitza jest następujące:

$$\begin{aligned}\sup_\pi\left(\mathbb{E}X_N^\pi - \theta\mathbb{D}X_N^2\right) &= \sup_\pi\left[x - \theta\left(\mathbb{E}\,(X_N^\pi)^2 - x^2\right)\right]\\ &= x + \theta x^2 - \theta\inf_\pi \mathbb{E}\,(X_N^\pi)^2.\end{aligned}$$

Dla problemu z funkcjonałem kosztu

$$J_N(x,\pi) = \mathbb{E}\,(X_N^\pi)^2$$

mamy $r(x) = x^2$,

$$\mathcal{A}r(x) = \inf_{b \in \mathbb{R}} \left(\frac{1}{2}(x+b)^2 + \frac{1}{2}(x-b)^2 \right) = x^2 = r(x),$$

gdzie minimum jest osiągalne dla $b = 0$. Stąd optymalną strategią jest nic nie obstawiać, a odpowiadający jej koszt wynosi x^2. W efekcie dla funcjonału Markowitza optymalną strategią jest nic nie obstawiać. Ponadto

$$\sup_{\pi} (\mathbb{E}\, X_N^{\pi} - \theta \mathbb{D}\, X_N^{\pi}) = X_0.$$

Załóżmy teraz, że $p \neq q$. Mamy $r(x) = r_1(x)$, gdzie

$$r_C(x) = C\,(x + \lambda/2)^2\,, \qquad C > 0$$

oraz

$$\mathcal{A}\phi(x) = \inf_{b \in \mathbb{R}} \left(p\phi(x+b) + q\phi(x-b) \right).$$

Stąd

$$\begin{aligned} \mathcal{A}\, r_C(x) &= C\,(x + \lambda/2)^2 + C \inf_{b \in \mathbb{R}} \left\{ 2(p-q)\,(x+\lambda/2)\,b + b^2 \right\} \\ &= C\left(1 - (p-q)^2\right)(x + \lambda/2)^2\,, \end{aligned}$$

gdzie infinum jest osiągalne dla

$$\widehat{b} = -(p-q)\,(x + \lambda/2)\,.$$

Tak więc z zasady indukcji wstecz Bellmana (Twierdzenie 2.2),

$$f(x,\lambda) = \inf_{\pi} J_N^{\lambda}(x,\pi) = \mathcal{A}^N r_1(x) = \left(1 - (p-q)^2\right)^N (x + \lambda/2)^2\,,$$

jest osiągalne dla strategii (stacjonarnej)

$$\widehat{\pi}^{\lambda} = \left(\widehat{b}_0^{\lambda}, \widehat{b}_1^{\lambda}, \ldots, \widehat{b}_{N-1}^{\lambda} \right)$$

gdzie

$$\widehat{b}_n^{\lambda}(x_0, \ldots x_n) = -(p-q)\,(x_n + \lambda/2)\,.$$

Dla zadanego $A \in \mathbb{R}$ szukamy teraz λ_A takiego, że

$$\mathbb{E}\, X_N^{\widehat{\pi}^{\lambda}} = A.$$

Przypomnijmy, że optymalna strategia dana jest wzorem

$$\widehat{b}_n^{\lambda}(X_n^{\widehat{\pi}^{\lambda}}) = -(p-q)\left(X_n^{\widehat{\pi}^{\lambda}} + \lambda/2 \right).$$

Mamy

$$
\begin{aligned}
\mathbb{E}\, X_N^{\widehat{\pi}^\lambda} &= \mathbb{E}\mathbb{E}\left(X_N^{\widehat{\pi}^\lambda} | X_{N-1}^{\widehat{\pi}^\lambda} \right) \\
&= \mathbb{E}\left\{ p\left[X_{N-1}^{\widehat{\pi}^\lambda} - (p-q)\left(X_{N-1}^{\widehat{\pi}^\lambda} + \lambda/2 \right) \right] + q\left[X_{N-1}^{\widehat{\pi}^\lambda} + (p-q)\left(X_{N-1}^{\widehat{\pi}^\lambda} + \lambda/2 \right) \right] \right\} \\
&= \left(1-(p-q)^2\right) \mathbb{E}\, X_{N-1}^{\widehat{\pi}^\lambda} - \frac{\lambda(p-q)^2}{2}.
\end{aligned}
$$

Stąd

$$
\begin{aligned}
\mathbb{E}\, X_N^{\widehat{\pi}^\lambda} &= \left(1-(p-q)^2\right)^N x - \sum_{k=0}^{N-1} \left(1-(p-q)^2\right)^k \frac{\lambda(p-q)^2}{2} \\
&= \left(1-(p-q)^2\right)^N x - \frac{\lambda}{2}(p-q)^2 \frac{1-\left(1-(p-q)^2\right)^N}{(p-q)^2} \\
&= \left(1-(p-q)^2\right)^N x - \frac{\lambda}{2}\left[1 - \left(1-(p-q)^2\right)^N \right].
\end{aligned}
$$

Niech

$$
\delta := \left(1-(p-q)^2\right)^N .
$$

Zauważmy, że $\delta \in (0,1)$. Mamy

$$
\mathbb{E}\, X_N^{\widehat{\pi}^\lambda} = \delta x - \frac{\lambda}{2}(1-\delta).
$$

Stąd

$$
\frac{\lambda}{2} = \frac{\delta x - \mathbb{E}\, X_N^{\widehat{\pi}^\lambda}}{1-\delta}.
$$

A więc $\mathbb{E}\, X_N^{\widehat{\pi}^\lambda} = A$ wtedy i tylko wtedy gdy

$$
\frac{\lambda}{2} = \frac{\delta x - A}{1-\delta}. \tag{2.20}
$$

W efekcie funkcja G występująca w punkcie (*iii*) ogólnego rozumowania wynosi

$$
G(A,x) = A + \theta\left[(A+\lambda/2)^2 - \delta\left(x+\lambda/2\right)^2 \right].
$$

Wstawiając do prawej strony $\lambda/2$ dane wzorem (2.20) otrzymujemy

$$
\begin{aligned}
G(A,x) &= A + \theta\left[\left(A + \frac{\delta x - A}{1-\delta} \right)^2 - \delta\left(x + \frac{\delta x - A}{1-\delta} \right)^2 \right] \\
&= A + \theta\left[\frac{\delta^2 (x-A)^2}{(1-\delta)^2} - \frac{\delta(x-A)^2}{(1-\delta)^2} \right] \\
&= A - \theta\frac{\delta}{1-\delta}(x-A)^2 \\
&= -\theta\frac{\delta}{1-\delta}A^2 + A\left(1 + \frac{2\theta\delta}{1-\delta}x\right) - \frac{\theta\delta}{1-\delta}x^2.
\end{aligned}
$$

Stąd (przypomnijmy, że $\delta \in (0,1)$) funkcja A osiąga maksimum dla

$$\widehat{A} := x + \frac{1-\delta}{2\theta\delta} = x + \frac{1 - \left(1-(p-q)^2\right)^N}{2\theta\left(1-(p-q)^2\right)^N}$$

Następnie

$$G(\widehat{A}, x) = x + \frac{1-\delta}{4\theta\delta} = x + \frac{1 - \left(1-(p-q)^2\right)^N}{4\theta\left(1-(p-q)^2\right)^N},$$

$$\frac{\lambda_{\widehat{A}}}{2} = \frac{\delta x - \widehat{A}}{1-\delta} = -x - \frac{1}{2\theta\delta} = -x - \frac{1}{2\theta\left(1-(p-q)^2\right)^N}.$$

Reasumując, dla problemu inwestora z funkcjonałem Markowitza maksymalny zysk wynosi

$$x + \frac{1 - \left(1-(p-q)^2\right)^N}{4\theta\left(1-(p-q)^2\right)^N},$$

a optymalną strategią jest nic nie obstawiać gdy $p = q = 1/2$, a gdy $p \neq q$ obstawiać w n-tej grze

$$\widehat{b}_n(X_0^{\widehat{\pi}}, X_1^{\widehat{\pi}}, \ldots, X_n^{\widehat{\pi}}) = -(p-q)\left(X_n^{\widehat{\pi}} - x - \frac{1}{2\theta\left(1-(p-q)^2\right)^N}\right).$$

Zauważmy, że optymalna strategia w chwili n zależy od kapitału gracza w chwili n jak i od kapitału początkowego x. Jest to zjawisko tak zwanego "time inconsistency".

2.11.2 Model inwestora z funkcjonałem Markowitza

Niech X_n będzie kapitałem inwestora w chwili n. Załóżmy, że jak w problemie Samuelsona w dowolnej chwili n inwestor może lokować kapitał B_n na rachunku ryzykownym z losową stopą zwrotu ξ_n a pozostały kapitał $X_n - B_n$ w banku na stałą stopę zwrotu $1 + r$. Przyjmujemy możliwości krótkiej sprzedarzy i krótkiego zakupu, a więc B_n lub $X_n - B_n$ mogą przyjmować wartości ujemne. Niech X_n^π oznacza kapitał inwestora w chwili n przy strategii alokacji kapitału π. Jak w problemie Samuelsona zakładamy, że (ξ_n) jest ciągiem niezależnych zmiennych losowych o tym samym rozkładzie. W przeciwieństwie do modelu Samuelsona inwestor nie konsumuje kapitału. Jego celem jest dla ustalonego skończonego horyzontu czasowego N zmaksymalizowanie

$$\mathbb{E}\, X_N^\pi - \theta \mathbb{D}\, X_N^\pi.$$

Dynamika kapitału dana jest wzorem

$$\begin{aligned} X_{n+1} &= (1+\xi_n)B_n + (1+r)(X_n - B_n) = (\xi_n - r)B_n + (1+r)X_n \\ &= (1+r)\left(\zeta_n B_n + X_n\right), \end{aligned}$$

gdzie $\zeta_n = (\xi_n - r)/(1+r)$, $n = 1, 2, \ldots$, jest ciągiem niezależnych zmiennych losowych o tym samym rozkładzie.

W pierwszym kroku rozważamy problem z kosztem

$$J_N^\lambda(x,\pi) = \mathbb{E}\left(X_N^\pi + \frac{\lambda}{2}\right)^2.$$

Mamy operator Bellmana

$$\mathcal{A}\psi(x) = \inf_{b\in\mathbb{R}} \mathbb{E}\,\psi\left((1+r)(\zeta_1 b + x)\right).$$

Niech

$$r(x) = \left(x + \frac{\lambda}{2}\right)^2.$$

Liczymy $\mathcal{A}r_{c,\lambda}(x)$ gdzie

$$r_{c,\lambda}(x) = c(x+\lambda/2)^2,$$

$$\begin{aligned}
\mathcal{A}r_{c,\lambda}(x) &= c\inf_{b\in\mathbb{R}} \mathbb{E}\left[(1+r)(\zeta_1 b + x) + \frac{\lambda}{2}\right]^2 \\
&= c(1+r)^2 \inf_{b\in\mathbb{R}} \mathbb{E}\left[\zeta_1 b + x + \frac{\lambda}{2(1+r)}\right]^2 \\
&= c(1+r)^2 \inf_{b\in\mathbb{R}} \left[b^2\mathbb{E}\,\zeta_1^2 + \left(x + \frac{\lambda}{2(1+r)}\right)^2 + b\mathbb{E}\,\zeta_1\left(2x + \frac{\lambda}{1+r}\right)\right].
\end{aligned}$$

Minimum jest osiągalne dla

$$b = -\frac{\mathbb{E}\,\zeta_1}{\mathbb{E}\,\zeta_1^2}\left[x + \frac{\lambda}{2(1+r)}\right].$$

Stąd

$$\begin{aligned}
\mathcal{A}r_{c,\lambda}(x) = c(1+r)^2\left[x + \frac{\lambda}{2(1+r)}\right]^2 &\left\{\frac{(\mathbb{E}\,\zeta_1)^2}{\mathbb{E}\,\zeta_1^2} + 1 - 2\frac{(\mathbb{E}\,\zeta_1)^2}{\mathbb{E}\,\zeta_1^2}\right\} \\
= c\left[1 - \frac{(\mathbb{E}\,\zeta_1)^2}{\mathbb{E}\,\zeta_1^2}\right](1+r)^2 &\left[x + \frac{\lambda}{2(1+r)}\right]^2 \\
= r_{\tilde{c},\tilde{\lambda}}(x),&
\end{aligned}$$

gdzie

$$\tilde{c} := c\left[1 - \frac{(\mathbb{E}\,\zeta_1)^2}{\mathbb{E}\,\zeta_1^2}\right](1+r)^2, \qquad \tilde{\lambda} = \frac{\lambda}{1+r}.$$

Stąd

$$\mathcal{A}^N r(x) = \left[1 - \frac{(\mathbb{E}\,\zeta_1)^2}{\mathbb{E}\,\zeta_1^2}\right]^N \left((1+r)^N x + \frac{\lambda}{2}\right)^2.$$

a optymalną strategią w chwili n jest

$$\widehat{B_n}(x) := -\frac{\mathbb{E}\,\zeta_1}{\mathbb{E}\,\zeta_1^2}\left[x + \frac{\lambda}{2(1+r)^{N-n}}\right].$$

Niech $\widehat{\pi}^\lambda$ będzie optymalną strategią. Liczymy teraz $\mathbb{E}\,X_N^{\widehat{\pi}^\lambda}$. Mamy

$$\begin{aligned}
\mathbb{E}\,X_n^{\widehat{\pi}^\lambda} &= (1+r)\left[\mathbb{E}\,X_{n-1}^{\widehat{\pi}^\lambda} + \mathbb{E}\,\mathbb{E}\left(\zeta_n \widehat{B}_{n-1}\left(X_{n-1}^{\widehat{\pi}^\lambda}\right) | X_{n-1}^{\widehat{\pi}^\lambda}\right)\right] \\
&= (1+r)\,\mathbb{E}\,X_{n-1}^{\widehat{\pi}^\lambda} - (1+r)\frac{(\mathbb{E}\,\zeta_1)^2}{\mathbb{E}\,\zeta_1^2}\left[\mathbb{E}\,X_{n-1}^{\widehat{\pi}^\lambda} + \frac{\lambda}{2(1+r)^{N-n+1}}\right] \\
&= (1+r)\left[1 - \frac{(\mathbb{E}\,\zeta_1)^2}{\mathbb{E}\,\zeta_1^2}\right]\mathbb{E}\,X_{n-1}^{\widehat{\pi}^\lambda} - \frac{(\mathbb{E}\,\zeta_1)^2}{\mathbb{E}\,\zeta_1^2}\frac{\lambda}{2(1+r)^{N-n}}.
\end{aligned}$$

Stąd

$$\mathbb{E}\,X_N^{\widehat{\pi}^\lambda} = (1+r)^N\left[1 - \frac{(\mathbb{E}\,\zeta_1)^2}{\mathbb{E}\,\zeta_1^2}\right]^N x - \frac{\lambda}{2(1+r)}C_N,$$

gdzie

$$\begin{aligned}
C_N &= \frac{(\mathbb{E}\,\zeta_1)^2}{\mathbb{E}\,\zeta_1^2}\sum_{k=0}^N \left[1 - \frac{(\mathbb{E}\,\zeta_1)^2}{\mathbb{E}\,\zeta_1^2}\right]^k \\
&= \frac{1 - \left[1 - \frac{(\mathbb{E}\,\zeta_1)^2}{\mathbb{E}\,\zeta_1^2}\right]^{N+1}}{\frac{(\mathbb{E}\,\zeta_1)^2}{\mathbb{E}\,\zeta_1^2}}\frac{(\mathbb{E}\,\zeta_1)^2}{\mathbb{E}\,\zeta_1^2} \\
&= 1 - \left[1 - \frac{(\mathbb{E}\,\zeta_1)^2}{\mathbb{E}\,\zeta_1^2}\right]^{N+1} > 0.
\end{aligned}$$

Niech

$$\rho := 1 - \frac{(\mathbb{E}\,\zeta_1)^2}{\mathbb{E}\,\zeta_1^2}.$$

Oczywiście $\rho \in (0,1)$. Mamy $C_N = 1 - \rho^{N+1}$. Z równania

$$A = \mathbb{E}\,X_N^{\widehat{\pi}^{\lambda_A}}$$

otrzymujemy, że

$$\frac{\lambda_A}{2} = \frac{1+r}{C_N}\left\{(1+r)^N \rho^N x - A\right\}.$$

Szukamy teraz $\widehat{A}$, które maksymalizuje funkcję

$$
\begin{aligned}
G(A,x) &= A + \theta\left(A + \frac{\lambda_A}{2}\right)^2 - \theta f(x,\lambda_A) \\
&= A + \theta\left(A + \frac{\lambda_A}{2}\right)^2 - \theta \mathcal{A}^N r(x) \\
&= A + \theta\left(A + \frac{\lambda_A}{2}\right)^2 - \theta\rho^N\left((1+r)^N x + \frac{\lambda_A}{2}\right)^2 \\
&= A + \theta\left[\left(1 - \frac{1+r}{C_N}\right)A + \frac{(1+r)^{N+1}}{C_N}\rho^N x\right]^2 \\
&\quad - \theta\rho^N\left\{-\frac{1+r}{C_N}A + \left[(1+r)^N + \frac{(1+r)^{N+1}}{C_N}\rho^N\right]x\right\} \\
&= a_N\theta A^2 + b_N A + d_N,
\end{aligned}
$$

gdzie

$$
\begin{aligned}
a_N &:= \left(1 - \frac{1+r}{C_N}\right)^2 - \rho^N\frac{(1+r)^2}{C_N^2}, \\
b_N &:= 1 + 2\theta\left(1 - \frac{1+r}{C_N}\right)\frac{(1+r)^{N+1}}{C_N}\rho^N x + 2\theta\rho^N\frac{1+r}{C_N}\left[(1+r)^N + \frac{(1+r)^{N+1}}{C_N}\rho^N\right]x \\
&= 1 + 2\theta x\rho^N\frac{(1+r)^{N+1}}{C_N}\left\{1 - \frac{1+r}{C_N} + \frac{1}{C_N} + \frac{1+r}{C_N}\frac{\rho^N}{C_N}\right\} \\
&= 1 + 2\theta x\rho^N\frac{(1+r)^{N+1}}{C_N}\left\{1 - \frac{r}{C_N} + \frac{1+r}{C_N}\frac{\rho^N}{C_N}\right\}, \\
d_N &:= \theta\frac{(1+r)^{2N+2}}{C_N^2}\rho^{2N}x^2 - \theta\rho^N\left[(1+r)^N + \frac{(1+r)^{N+1}}{C_N}\rho^N\right]^2 x^2 \\
&= \theta x^2\rho^N(1+r)^{2N}\left\{\frac{(1+r)^2}{C_N^2}\rho^N - \left[1 + \frac{1+r}{C_N}\rho^N\right]^2\right\}.
\end{aligned}
$$

Stąd

$$
a_N = \frac{1}{C_N^2}\left\{(C_N - (1+r))^2 - \rho^N(1+r)^2\right\} = \frac{(1+r)^2}{C_N^2}\left\{\left[\frac{1-\rho^{N+1}}{1+r} - 1\right]^2 - \rho^N\right\}.
$$

A więc $a_N < 0$ wtedy i tylko wtedy gdy

$$
\frac{\rho^{N/2} - \rho^{N+1}}{1 - \rho^{N/2}} > r. \tag{2.21}
$$

Załóżmy, że zachodzi (2.21). Wówczas dla ustalonego x funkcja $G(A,x)$ osiąga maksimum dla

$$
\widehat{A} = -\frac{b_N}{2a_N\theta}.
$$

Ponadto

$$G(\widehat{A}, x) = -\frac{b_N^2}{4a_N\theta} + d_N,$$

o optymalna strategia dana jest wzorem

$$\widehat{B_n}(X_n^{\widehat{\pi}}) := -\frac{\mathbb{E}\,\zeta_1}{\mathbb{E}\,\zeta_1^2}\left\{X_n^{\widehat{\pi}} + \frac{1}{2(1+r)^{N-n}}\left[\frac{1+r}{C_N}\left((1+r)^N\rho^N x + \frac{b_N}{2a_N\theta}\right)\right]\right\}.$$

Gdy (2.21) nie zachodzi wówczas supremum funkcjonału Markowitza po wszystkich strategiach dopuszczalnych jest równe $+\infty$, ale nie jest osiągalne dla żadnej strategii.

3

Sterowanie na nieskończonym przedziale czasowym

Naszym celem jest znajdowanie optymalnej strategii dla funkcjonału

$$J_\infty(\pi, X_0) = \mathbb{E} \sum_{n=0}^{+\infty} \gamma^n q\left(X_n^\pi, u_n(X_0^\pi, \dots, X_n^\pi)\right),$$

gdzie $q\colon E \times U \to [0, +\infty)$ jest funkcją mierzalną, proces $\{X_n^\pi\}$ zadany jest rekurencyjnie,

$$X_{n+1}^\pi = F\left(X_n^\pi, u_n(X_0^\pi, \dots, X_n^\pi), \xi_{n+1}\right),$$

$u_n\colon E^{n+1} \to U$, $n \geq 0$ są odwzorowaniami mierzalnymi, takim że, patrz Rozdział 2,

$$u_n(x_0, \dots, x_n) \in U(x_n), \qquad \forall\, n,\ \forall\, x_n.$$

Oczywiście w przypadku gdy J_∞ jest funkcjonałem zysku to staramy się znaleźć strategie $\widehat{\pi}$, która, przy zadanej wartości początkowej X_0, go maksymalizuje. Gdy, J_∞ jest funkcjonałem kosztu to szukamy strategii, która go minimalizuje.

3.1 Zasadnicze twierdzenia

Jak w Rozdziale 2, wprowadzimy operator $\mathcal{A}$ w następujący sposób:

$$\mathcal{A}v(x) = \sup_{u \in U(x)} \left(q(x,u) + \gamma P^u v(x)\right),$$

gdy J_∞ jest zyskiem, oraz

$$\mathcal{A}v(x) = \inf_{u \in U(x)} \left(q(x,u) + \gamma P^u v(x)\right),$$

gdy J_∞ jest kosztem. Przypomnijmy, że

$$P^u h(x) := \mathbb{E}\, h(F(x, u, \xi_0)).$$

Zawuważmy najpierw, że operator $\mathcal{A}$ jest *monotoniczny*, to znaczy dla dowolnych funkcji mierzalnych nieujemnych $v_1, v_2 \colon E \to [0, +\infty]$ mamy

$$v_1(x) \le v_2(x),\ \forall\, x \in E \quad \Longrightarrow \quad \mathcal{A}v_1(x) \le \mathcal{A}v_2(x),\ \forall\, x \in E.$$

Będziemy zakładali, że ciąg iteracji $\mathcal{A}^n(0)$, $n = 0, 1, \ldots$, gdzie 0 oznacza funkcje stałą 0, jest ciągiem dobrze określonych funkcji mierzalnych. Ponieważ $\mathcal{A}$ jest monotoniczny mamy

$$\forall\, n \le m,\ \forall\, x \in E, \qquad 0 \le \mathcal{A}^n(0)(x) \le \mathcal{A}^m(0)(x).$$

Stąd istnieje granica punktowa

$$v_\infty(x) = \lim_{n \to +\infty} \mathcal{A}^n(0)(x), \qquad x \in E.$$

Oczywiście v_∞ może przyjmować wartość $+\infty$. Zasadniczymi wynikami są dwa następujące twierdzenia. Pierwsze dotyczy przypadku funkcjonału kosztu.

Twierdzenie 3.1 *Załóżmy, że J_∞ jest kosztem. Wówczas dla dowolnego $x \in E$ i dla dowolnej strategi π mamy*

$$v_\infty(x) \le \mathcal{A}v_\infty(x), \qquad v_\infty(x) \le J_\infty(\pi, x).$$

Ponadto, jeżeli dla wszystkich $x \in E$,

$$v_\infty(x) = \mathcal{A}v_\infty(x) = q(x, u_\infty(x)) + \gamma P^{u_\infty(x)} v_\infty(x),$$

to strategia

$$\pi_\infty = (u_\infty(x_0), u_\infty(x_1), \ldots)$$

jest optymalna.

Dowód Ponieważ $\mathcal{A}^n(0) \le v_\infty$ z monotoniczności $\mathcal{A}$ mamy $\mathcal{A}^{n+1}(0) \le \mathcal{A}v_\infty$. Stąd $v_\infty \le \mathcal{A}v_\infty$. Z twierdzenia Bellmana dla skończonego horyzontu, dla dowolnej strategi π i dla dowolnych N i punktu startu $X_0 = x \in E$ mamy

$$\mathbb{E} \sum_{n=0}^{N-1} \gamma^n q\left(X_n^\pi, u_n(X_0^\pi, \ldots, X_n^\pi)\right) \ge \mathcal{A}^N(0)(x).$$

Stąd $v_\infty(x) \le J_\infty(\pi, x)$, $x \in E$. Aby wykazać drugą część twierdzenia zauważmy, że

$$\begin{aligned} J_N(\pi_\infty, x) &= \mathbb{E} \sum_{n=0}^{N-1} \gamma^n q\left(X_n^{\pi_\infty}, u_\infty(X_n^{\pi_\infty})\right) \\ &\le \mathbb{E}\left(\sum_{n=0}^{N-1} \gamma^n q\left(X_n^{\pi_\infty}, u_\infty(X_n^{\pi_\infty})\right) + \gamma^N v_\infty(X_N^{\pi_\infty}) \right) \\ &= v_\infty(x), \end{aligned}$$

bo $\mathcal{A}v_\infty = v_\infty$. $\square$

Twierdzenie 3.2 *Załóżmy, że J_∞ jest zyskiem. Wówczas dla dowolnego $x \in E$ i dla dowolnej strategii π mamy*

$$v_\infty(x) = \mathcal{A}v_\infty(x), \qquad v_\infty(x) \geq J_\infty(\pi, x).$$

Ponadto, jeżeli dla wszystkich $x \in E$,

$$v_\infty(x) = \mathcal{A}v_\infty(x) = q(x, u_\infty(x)) + \gamma P^{u_\infty(x)} v_\infty(x),$$

to strategia

$$\pi_\infty = (u_\infty(x_0), u_\infty(x_1), \ldots)$$

jest optymalna, gdy spełniony jest dodatkowy warunek

$$\lim_{N\to+\infty} \mathbb{E}\,\gamma^N v_\infty(X_N^{\pi_\infty}) = 0. \tag{3.1}$$

Dowód Pokażemy, że $\mathcal{A}v_\infty = v_\infty$. Zauważmy, że

$$\begin{aligned}\mathcal{A}^{n+1}(0)(x) &= \sup_{u\in U(x)} \left(q(x,u) + \gamma P^u \mathcal{A}^n(0)(x)\right) \\ &\geq q(x,u) + \gamma P^u \mathcal{A}^n(0)(x), \qquad \forall\, x \in E,\ \forall\, u \in U(x).\end{aligned}$$

Stąd, po przejściu do granicy otrzymujemy

$$v_\infty(x) \geq q(x,u) + \gamma P^u v_\infty(x), \qquad \forall\, x \in E,\ \forall\, u \in U(x).$$

Stąd, biorąc supremum po u otrzymujemy $\mathcal{A}v_\infty \leq v_\infty$. Ponieważ z monotoniczności $\mathcal{A}$, $\mathcal{A}v_\infty \geq v_\infty$, otrzymujemy $\mathcal{A}v_\infty = v_\infty$.

Pokażemy teraz, że dla dowolnej strategii π, $v_\infty(x) \geq J_\infty(\pi, x)$. Wynika to jednak wprost z twierdzenia Bellmana dla skończonego horyzontu. Istotnie dowolnych N i punktu startu $X_0 = x \in E$ mamy

$$\mathbb{E} \sum_{n=0}^{N-1} \gamma^n q\left(X_n^\pi, u_n(X_0^\pi, \ldots, X_n^\pi)\right) \leq \mathcal{A}^N(0)(x) \leq v_\infty(x).$$

Optymalność π_∞ przy założeniu (3.1) będzie natychmiastowym wnioskiem z równości

$$J_\infty(\pi_\infty, x) = v_\infty(x), \qquad x \in E.$$

Aby ją wykazać połóżmy

$$\tilde{J}_N(\pi, x) := \mathbb{E}\left(\sum_{n=0}^{N-1} \gamma^n q(X_n, u_n) + \gamma^N v_\infty(X_N)\right).$$

Korzystając z założenia otrzymujemy

$$J_\infty(\pi_\infty, x) = \lim_{N\to+\infty} \tilde{J}_N(\pi_\infty, x).$$

Z twierdzenia Bellmana i z faktu, że $\mathcal{A}v_\infty = v_\infty$,

$$\tilde{J}_N(\pi_\infty, x) = v_\infty(x).$$

Stąd wynika żądana konkluzja. $\square$

3.2 Problemy inwestora

Znajdziemy rozwiązanie problemów inwestora na nieskończonym przedziale czasowym z funkcjami satysfakcji $U(c) = c^\alpha$, $\alpha \in (0,1)$ i $U(c) = \log c$. Problemy inwestora ze skończonym horyzontem czasowym rozważane były w Rozdziale 2.7.

3.2.1 Klasyczny problem Samuelsona; $U(c) = c^\alpha$

W tym przypadku funkcjonał zysku (satysfakcji) na postać

$$J_\infty(\pi, X_0) = \mathbb{E} \sum_{n=0}^{+\infty} \gamma^n c_n^\alpha.$$

Niech $r_z(x) = zx^\alpha$. Przypomnijmy, patrz Rozdział 2.7.1, że operator Bellmana spełnia

$$\mathcal{A} r_z = \widehat{f}(z) r_1,$$

gdzie

$$\widehat{f}(z) := \left(1 + (z\gamma\widehat{\lambda})^{1/(1-\alpha)}\right)^{1-\alpha}$$

a

$$\widehat{\lambda} := \sup_{b\in[0,1]} \mathbb{E}\left((1+r)b + (1-b)(1+\xi_0)\right)^\alpha,$$

Ponadto niech

$$\widehat{g}(z) := \frac{1}{1 + (z\gamma\widehat{\lambda})^{1/(1-\alpha)}},$$

oraz niech

$$\widehat{b} := \operatorname{Arg} \sup_{b\in[0,1]} \mathbb{E}\left((1+r)b + (1-b)(1+\xi_0)\right)^\alpha.$$

Wówczas, patrz Rozdział 2.7.1,

$$(g(z)x, \widehat{b}) = \operatorname{Arg} \sup_{c\in[0,x],\, b\in[0,1]} \left\{ c^\alpha + \gamma P^{(c,b)} r_z(x) \right\}.$$

Stąd, dla $r_0 = 0$ otrzymujemy

$$\mathcal{A}^n(0)(x) = \widehat{f}^n(0) x^\alpha, \qquad x \in [0,+\infty),\ n \in \mathbb{N}.$$

Twierdzenie poniżej jest więc bezpośrednim wnioskiem z Twierdzenia 3.1.

Twierdzenie 3.3 *Jeżeli $\gamma\widehat{\lambda} < 1$ to ciąg $\{\widehat{f}^n(\omega)\}$ zbiega do skończonej granicy*

$$\widehat{\omega} = \frac{1}{\left(1 - (\gamma\widehat{\lambda})^{1/(1-\alpha)}\right)^{1/(1-\alpha)}},$$

$v_\infty(x) = \widehat{\omega} x^\alpha$, $x \geq 0$, *oraz optymalna strategia*

$$\pi_\infty = (u_\infty(x_0), u_\infty(x_1), \ldots)$$

dana jest wzorem

$$u_\infty(x) = \left(c_\infty x, \widehat{b}\right), \qquad c_\infty = \widehat{g}(\widehat{\omega}) = \frac{1}{1 + (\widehat{\omega}\gamma\widehat{\lambda})^{1/(1-\alpha)}}.$$

Gdy $\gamma\widehat{\lambda} \geq 1$ *to* $v_\infty = +\infty$ *i istnieje strategia, dla której zysk jest nieskończony.*

3.2.2 Przypadek logarytmicznej funkcji satysfakcji; $U(c) = \log c$

W Rozdziale 2.7.2 pokazaliśmy, że dla funkcji

$$r_{z,v}(x) = z + v \log x,$$

mamy

$$\mathcal{A} r_{z,v} = r_{f(z,v),g(v)}$$

gdzie

$$f(z,v) = \gamma(z + v\widehat{\lambda}) + \gamma v \log \gamma v - (1+\gamma v)\log(1+\gamma v),$$
$$g(v) = \frac{1}{1+\gamma v},$$

a

$$\widehat{\lambda} := \sup_{b \in [0,1]} \mathbb{E} \log\left(1 + \xi_1 + b(r - \xi_1)\right).$$

Załóżmy, że supremum jest osiągane w $\widehat{b}$. Wówczas

$$(g(v)x, \widehat{b}) = \operatorname{Arg} \sup_{c \in [0,x],\, b \in [0,1]} \left\{ \log c + \gamma P^{(c,b)} r_{z,v}(x) \right\}.$$

Ponieważ $0 = r_{0,0}$ więc

$$v_\infty(x) = r_{\widehat{z},\widehat{v}}(x),$$

gdzie $\widehat{z}$ i $\widehat{v}$ są nieujemnymi skończonymi rozwiązaniami układu równań

$$f(\widehat{z},\widehat{v}) = \widehat{z}, \qquad g(\widehat{v}) = \widehat{v}.$$

Mamy

$$\widehat{v} = \frac{1}{1+\gamma\widehat{v}},$$

więc

$$\gamma\widehat{v}^2 + \widehat{v} - 1 = 0.$$

Stąd

$$\widehat{v} = \frac{-1 + \sqrt{1+4\gamma}}{2\gamma}.$$

Następnie

$$\gamma(\widehat{z}+\widehat{v}\widehat{\lambda})+\gamma\widehat{v}\log\gamma\widehat{v}-(1+\gamma\widehat{v})\log(1+\gamma\widehat{v})=\widehat{z}.$$

Czyli

$$\widehat{z}=\frac{1}{1-\gamma}\left[\gamma\widehat{v}\widehat{\lambda}+\gamma\widehat{v}\log\gamma\widehat{v}-(1+\gamma\widehat{v})\log(1+\gamma\widehat{v})\right].$$

Oczywiście γ i $\widehat{\lambda}$ powinny być takie, że $\widehat{z}>0$. W przeciwnym razie $v_\infty\equiv+\infty$. W przypadku skończonej v_∞, optymalnym podziałem środków jest $\widehat{b}$, a optymalną konsumpcją jest

$$\widehat{c}_n(x_n)=g(\widehat{v})x_n.$$

3.3 Model z losową stopą krótką

Niech V_n oznacza kapitał inwestora w chwili n. W każdej chwili inwestor konsumuje $C_n\le V_n$ a resztę V_n-C_n lokuje na rachunku bankowym. Załóżmy, że stopa R_n ofereowana w banku w chwili n na okres od n do $n+1$, jest łańcuchem Markowa o wartościach w zbiorze $D\subseteq[0,+\infty)$. Niech P będzie operatorem przejścia łańcucha (R_n). Dynamika kapitału jest następująca

$$V_{n+1}=R_n(V_n-C_n).$$

Celem inwestora jest znalezienie konsumpcji $\pi=(C_n)$ która maksymalizuje jego satysfakcje

$$J(R_0,V_0,\pi)=\mathbb{E}\sum_{n=0}^{\infty}\gamma^{-n}C_n^{\alpha}. \tag{3.2}$$

W (16.3), $\gamma\in(0,1)$ opisuje stopę inflacji a $\alpha\in(0,1)$. Niech V^π będzie kapitałem inwestora odpowiadającym strategi konsumpcji $\pi=(C_n)$. Sterowanym łańcuchem Markowa jest $X^\pi=(V^\pi,R)$.

Operator Bellmana dla naszego problemu ma postać

$$\begin{aligned}\mathcal{A}f(R,v)&=\sup_{0\le c\le v}\left\{c^\alpha+\gamma\mathbb{E}\left(f(R_{n+1},V_{n+1})|R_n=R,V_n=v,C_n=c\right)\right\}\\&=\sup_{0\le c\le v}\left\{c^\alpha+\gamma\mathbb{E}\left(f(R_{n+1},R(v-c))|R_n=R\right)\right\}\\&=\sup_{0\le c\le v}\left\{c^\alpha+\gamma Pf(\cdot,R(v-c))(R)\right\}.\end{aligned}$$

Niech $\mathcal{A}^0$ będzie operatorem identyczności a $\mathcal{A}^n$ niech będzie n-tą iteracją operatora $\mathcal{A}$. Niech

$$W_n(R,v):=\mathcal{A}^n(0)(R,v),\qquad n=0,1,\ldots,\ R\in D,\ v\ge 0.$$

Oczywiście
$$\mathcal{A}(0)(R,v) = \sup_{0\le c\le v} c^\alpha = v^\alpha,$$
gdzie supremum jest osiągalne dla $c = u_1(v,R) = v$. W celu obliczenia iteracji $\mathcal{A}$ potrzebujemy następującego elementarnego lematu.

Lemma 3.1. *Dla dowolnego $d \ge 0$,*
$$\sup_{0\le c\le v} (c^\alpha + d(v-c)^\alpha) = v^\alpha h(d)$$
oraz supremum jest osiągalne dla $c = g(d)v$, gdzie
$$h(d) := \frac{d^{\alpha/(\alpha-1)} + d}{\left(1 + d^{1/(\alpha-1)}\right)^\alpha} = \frac{1 + d^{1/(1-\alpha)}}{\left(1 + d^{1/(1-\alpha)}\right)^\alpha} = \left(1 + d^{1/(1-\alpha)}\right)^{1-\alpha} \tag{3.3}$$
oraz
$$g(d) := \frac{d^{1/(\alpha-1)}}{1 + d^{1/(\alpha-1)}} = \left(1 + d^{1/(1-\alpha)}\right)^{-1}. \tag{3.4}$$

Wracając do obliczenia iteracji $\mathcal{A}$. Mamy
$$\begin{aligned}\mathcal{A}^2(0)(r,v) &= \sup_{0\le c\le v} (c^\alpha + \gamma(r(v-c))^\alpha)\\ &= \sup_{0\le c\le v} (c^\alpha + \gamma r^\alpha (v-c)^\alpha)\\ &= v^\alpha h(\gamma r^\alpha) = v^\alpha \phi(r),\end{aligned}$$
gdzie
$$\phi(r) := h(\gamma r^\alpha) = \left(1 + (\gamma r^\alpha)^{1/(1-\alpha)}\right)^{1-\alpha}, \qquad r \ge 0. \tag{3.5}$$
Ponadto, supremum jest osiągalne w
$$u_2(r,v) = vg(\gamma r^\alpha).$$
Zauważmy, że ϕ jest rosnąca gdy h jest rosnąca. Następnie
$$\begin{aligned}\mathcal{A}^3(0)(r,v) &= \sup_{0\le c\le v} (c^\alpha + \gamma r^\alpha (v-c)^\alpha\, P\phi(r))\\ &= \sup_{0\le c\le v} (c^\alpha + \gamma\, r^\alpha P\phi(r)\,(v-c)^\alpha)\\ &= v^\alpha h\left(\gamma\, r^\alpha\, P\phi(r)\right) = v^\alpha \phi\left(r\,(P\phi(r))^{1/\alpha}\right)\end{aligned}$$
i supremum jest osiągalne dla
$$u_3(r,v) := vg\left(\gamma\, r^\alpha P\phi(r)\right).$$
Stąd mamy następujący wynik.

Lemma 3.2. *Dla* $n = 0, 1, \ldots,$

$$\mathcal{A}^n(0)(r,v) = v^\alpha \psi_{n-1}(r), \qquad r \in D,\ v \geq 0,$$

gdzie $\psi_{-1} \equiv 0$ *oraz*

$$\psi_n(r) = \phi\left(r\,(P\psi_{n-1}(r))^{1/\alpha}\right), \qquad n = 0, 1, \ldots,\ r \in D,$$

a ϕ *jest zdefiniowane jak w* (3.5). *Ponadto*

$$\mathcal{A}^n(0)(r,v) = (c_n(r,v))^\alpha + \gamma P \mathcal{A}^{n-1}(0)\left(\cdot, r(v - c_n(r,v))\right)(r),$$

gdzie

$$c_n(r,v) := g\left(\gamma\, r^\alpha P\psi_{n-1}(r)\right) v, \qquad r \in D,\ v \geq 0$$

i g *jest zdefiniowane jak w* (3.4).

Przypomnijmy, że v_∞ jest granicą v_n gdy $n \uparrow \infty$. Z Lematu 3.2,

$$v_n(r,v) = v^\alpha \psi_{n-1}(r), \qquad r \in D,\ v \geq 0.$$

Zauważmy, że ϕ jest rosnąca i P jest monotoniczne; to znaczy, że $P\phi_1 \leq P\phi_2$ jeżeli $\phi_1 \leq \phi_2$. Stąd (ψ_n) jest ciągiem rosnącym. Tak więc zbiega punktowo do mierzalnej nieujemnej funkcji ψ (o wartościach w $[0, +\infty]$). Ponieważ ϕ jest ciągła i rosnąca granica ψ spełnia równanie funkcyjne

$$\varphi(r) = \phi\left(rP\varphi(r)\right) = \left(1 + (\gamma\, r^\alpha\, P\varphi(r))^{1/(1-\alpha)}\right)^{1-\alpha}, \qquad r \in D. \tag{3.6}$$

Oczywiście $\varphi \equiv +\infty$ jest rozwiązaniem (3.6). Zauważmy, że ψ jest jedynym minimalnym nieujemnym rozwiązaniem (3.6). Istotnie, załóżmy, że φ jset mierzalnym nieujemnym rozwiązaniem (3.6). Wówczas $0 \equiv \psi_{-1} \leq \varphi$. Załóżmy, że $\psi_{n-1}(r) \leq \varphi(r)$ dla $r \in D$. Z monotoniczności ϕ oraz P,

$$\psi_n(r) = \phi\left(r\,(P\psi_{n-1}(r))^{1/\alpha}\right) \leq \phi\left(r\,(P\varphi(r))^{1/\alpha}\right) = \varphi(r), \qquad \forall\, r \in D.$$

Stąd, stosując zasadę indukcji, $\psi_n(r) \leq \varphi(r)$ dla wszystkich $r \in D$, a więc $\psi \leq \varphi$. Zauważ, że g dane przez (3.4) jest funkcją ciągłą a więc ponieważ (ψ_n) jest zbieżny $P\psi_n$ zbiega do $P\psi$.

Twierdzenie 3.4 (*i*) *Granica* v_∞ *jest dana przez*

$$v_\infty(r,v) = v^\alpha \psi(r), \qquad r \in D,\ v \geq 0,$$

gdzie ψ *jest jedynym minimalnym nieujemnym rozwiązaniem równania funkcyjnego* (3.6).
(*ii*) *Załóżmy, że* (R_n) *przyjmuje wartości w przedziale* $D = [0, M]$, *gdzie* $M < \infty$. *Jeżeli*

$$\gamma M^\alpha < 1, \tag{3.7}$$

to minimalne nieujemne rozwiązanie ψ równania (3.6) *jest ograniczoną funkcją na D. Ponadto optymalna konsumpcja c_∞ dana jest wzorem*

$$\begin{aligned} c_\infty(r,v) &= g\left(\gamma\, r^\alpha P\psi(r)\right) v \\ &= \left(1 + \left(\gamma\, r^\alpha P\psi(r)\right)^{1/(1-\alpha)}\right)^{-1} v, \quad r \in D,\, v \geq 0. \end{aligned}$$

Dowód. Pierwsza część wynika wprost z Lematu 3.2. Aby pokazać drugą część musimy dowieść, że granica punktowa

$$\psi(r) := \lim_{n\to\infty} \psi_n(r), \qquad r \in D,$$

jest skończona. Przypomnijmy, że ciąg $\{\psi_n(r)\}$ jest rosnący dla każdego $r \in D$.

Niech $B_{b,+} := B_b(D, [0,\infty))$ oznacza przestrzeń mierzalnych nieujemnych funkcji na D, oraz niech

$$\|\varphi\|_\infty := \sup_{r\in[0,M]} \varphi(r), \qquad \varphi \in B_{b,+}.$$

Mamy

$$\begin{aligned} 0 \leq \psi_n(r) &= \left(1 + \left(\gamma\, r^\alpha P\psi_{n-1}(r)\right)^{1/(1-\alpha)}\right)^{1-\alpha} \\ &\leq 1 + \gamma r^\alpha P\psi_{n-1}(r) \\ &\leq 1 + \gamma\, M^\alpha\, \|\psi_{n-1}\|_\infty. \end{aligned}$$

Tak więc

$$\|\psi_n\|_\infty \leq 1 + \gamma\, M^\alpha \|\psi_{n-1}\|_\infty, \qquad \forall\, n = 0,1,\ldots.$$

Biorąc pod uwagę, że $\psi_{-1} \equiv 0$ oraz $\psi_0 \equiv 1$ otrzymujemy

$$\|\psi_n\|_\infty \leq \sum_{k=0}^{n} (\gamma M^\alpha)^k \leq \frac{1}{1-\gamma M^\alpha} < \infty, \quad \forall\, n = 0,1,\ldots.$$

3.4 Przykład

Załóżmy, że (r_n) jest łańcuchem Markowa o wartościach w skończonym zbiorze $D = \{a_1, a_2, \ldots, a_m\}$, gdzie $0 \leq a_1 < a_2 < \ldots < m$, z macierzą przejścia $p_{i,j} := \mathbb{P}\left(R_{n+1} = a_j | R_n = a_i\right)$. Wówczas

$$Pf(a_i) = \sum_{j=1}^{m} p_{i,j} f(a_j), \qquad i = 1,\ldots,m.$$

Stąd równanie (3.6) ma postać

$$\varphi(a_i) = \left(1 + \left(\gamma\, a_i^\alpha \sum_{j=1}^{n} p_{i,j}\varphi(a_j)\right)^{1/(1-\alpha)}\right)^{1-\alpha}, \ i = 1,\ldots,n. \tag{3.8}$$

3.5 Przypadek ryzykownych akcji

W przypadku rachónku bankowego jest racjonalne założyć, że w chwili n znamy stopę oferowaną przez bank na odcinek $[n, n+1]$. W przypadku rachónku akcyjnego racjonalniej jest założyć, że dynamika kapitału jest dana wzorem

$$V_{n+1} = R_{n+1}(V_n - C_n).$$

Jak w przypadku inwestycji na rachunku bankowym inwestor szuka strategii maksymalizującej satysfakcję

$$J((C_n), R_0, V_0) := \mathbb{E}\sum_{k=0}^{\infty} \gamma^n C_n^{\alpha}.$$

Odpowiadający operator Bellmana ma postać

$$\begin{aligned}\mathcal{A}f(r,v) &= \sup_{0\le c\le v} \left\{c^{\alpha} + \gamma\mathbb{E}\left(f(R_{n+1}, V_{n+1})|R_n = r, V_n = v, C_n = c\right)\right\}\\ &= \sup_{0\le c\le v} \left\{c^{\alpha} + \gamma Pf(\cdot, \cdot(v-c))(r)\right\}.\end{aligned}$$

Niech

$$v_n(r,v) = \mathcal{A}^n(0)(r,v).$$

Oczywiście $v_1(r,v) = v^{\alpha}$. Następnie, z Lematu 3.1,

$$\mathcal{A}^2(0)(r,v) = \sup_{0\le c\le v} \left(c^{\alpha} + \gamma(v-c)^{\alpha} Pr^{\alpha}\right) = v^{\alpha} h\left(\gamma Pr^{\alpha}\right),$$

$$\mathcal{A}^3(0)(r,v) = v^{\alpha} h\left(\gamma P\left(r^{\alpha} h(\gamma Pr^{\alpha})\right)\right).$$

Stąd mamy następujący wynik.

Lemat 3.1 *Dla* $n = 0, 1, \ldots,$

$$\mathcal{A}^n(0)(r,v) = v^{\alpha}\psi_{n-1}(r), \qquad r \in D,\ v \ge 0,$$

gdzie $\psi_{-1} \equiv 0$ *oraz*

$$\psi_n(r) = h\left(\gamma\, P\left(r^{\alpha}\psi_{n-1}(r)\right)\right), \qquad n = 0, 1, \ldots,\ r \in D,$$

a h *jest zdefiniowane jak w* (3.3)*. Ponadto,*

$$\mathcal{A}^n(0)(r,v) = \left(c_n(r,v)\right)^{\alpha} + \gamma P\mathcal{A}^{n-1}(0)\left(\cdot, \cdot(v - c_n(\cdot, v))\right)(r),$$

gdzie

$$c_n(r,v) := g\left(\gamma\, P\, r^{\alpha}\psi_{n-1}(r)\right) v, \qquad r \in D,\ v \ge 0$$

oraz g *jest zdefiniowane jak* (3.4)*.*

Niech

$$\psi(r) = \lim_{n\to\infty} \psi_n(r), \qquad r \in D.$$

Wówczas ψ jest minimalnym nieujemnym rozwiązaniem równania funkcyjnego

$$\varphi(r) = h\left(\gamma P\left(r^\alpha \psi(r)\right)\right), \qquad r \in D. \tag{3.9}$$

Twierdzenie 3.5 *(i) Granica v_∞ jest dana przez*

$$v_\infty(r, v) = v^\alpha \psi(r), \qquad r \in D,\ v \geq 0,$$

gdzie ψ jest (jedynym) minimalnym nieujemnym rozwiązaniem równania (3.9).
(ii) Załóżmy, że (R_n) przyjmuje wartości w odcinku $D = [0, M]$, gdzie $M < \infty$. Jeżeli

$$\gamma M^\alpha < 1, \tag{3.10}$$

wtedy minimalne nieujemne rozwiązanie ψ równania (3.9) *jest funkcją ograniczoną na D. Ponadto optymalną konsumpcja c_∞ jest dana wzorem*

$$\begin{aligned} c_\infty(r, v) &= g\left(\gamma\, P r^\alpha \psi(r)\right) v \\ &= \left(1 + \left(\gamma\, P r^\alpha \psi(r)\right)^{1/(1-\alpha)}\right)^{-1} v, \quad r \in D,\, v \geq 0. \end{aligned}$$

Dowód. Pierwsza część wynika z Lematu 3.1. Aby pokazać drugą część zauważmy, że

$$\begin{aligned} 0 \leq \psi_n(r) &= \left(1 + \left(\gamma\, P r^\alpha \psi_{n-1}(r)\right)^{1/(1-\alpha)}\right)^{1-\alpha} \\ &\leq 1 + \gamma M^\alpha P \psi_{n-1}(r) \\ &\leq 1 + \gamma\, M^\alpha \left\|\psi_{n-1}\right\|_\infty, \end{aligned}$$

i wnioskujemy jak w dowodzie Twierdzenia 3.4.

Przyklad 3.1 Załóżmy, że (r_n) jest łańcuchem Markowa w skończonej przestrzeni stanów $D = \{a_1, a_2, \ldots, a_m\}$, gdzie $0 \leq a_1 < a_2 < \ldots < m$, z macierzą przejścia $p_{i,j} := \mathbb{P}\left(R_{n+1} = a_j | R_n = a_i\right)$. Wówczas równanie (3.9) przyjmuje postać

$$\varphi(a_i) = \left(1 + \left(\gamma \sum_{j=1}^{n} p_{i,j} a_j^\alpha \varphi(a_j)\right)^{1/(1-\alpha)}\right)^{1-\alpha}, \ i = 1, \ldots, n. \tag{3.11}$$

3.6 Model wieloskładnikowy

Załóżmy, że inwestor może lokować swój kapitał kupując l-różnych akcji oraz lokując na rachunku bankowym. Niech $R_{n,k}$ będzie stopą zwrotu w chwili n z

k-tej akcji. Niech $R_n = (R_{n,0}, \ldots, R_{n,l})$. Załóżmy, że $R = (R_n)$ jest łańcuchem Markowa o wartościach w zbiorze $D \subset [0, +\infty)^{k+1}$. Niech $R_{n,0}$ będzie stopą oferowaną przez bank. Niech V_n będzie kapitałem inwestora oraz niech $0 \leq C_n \leq V_n$ będzie konsumpcją w chwili n. Przez $\theta_n = (\theta_{n,0}, \ldots, \theta_{n,m})$ oznaczamy strategię inwestowania w chwili n. Niech

$$\varDelta_m := \left\{ (\theta_0, \ldots, \theta_m) \colon \theta_k \geq 0,\ \forall\, k, \quad \sum_{k=0}^{m} \theta_k = 1 \right\}.$$

Zakładamy, że $\Theta_n \in \varDelta_m$ for $n = 0, 1, \ldots$.

Dynamika kapitału odpowiadająca strategii $\pi = (\theta_n, C_n)$ jest dana wzorem

$$V_{n+1} = \theta_0 R_{n,0}(V_n - C_n) + \sum_{k=1}^{m} \theta_{n,k} R_{n+1,k} \, (V_n - C_n) \, .$$

Celem inwestora jest maksymalizacja oczekiwanej satysfakcji

$$J(V_0, R_0, \pi) = \mathbb{E} \sum_{n=0}^{\infty} \gamma^n C_n^{\alpha},$$

gdzie $\gamma \in (0, 1]$ jest dyskontem.

Niech P będzie operatorem przejścia dla łańcucha R. Operator Bellmana dany jest wzorem

$$\mathcal{A} f(r, v) = \sup_{0 \leq c \leq v,\ \theta \in \varDelta_m} \left\{ c^{\alpha} + \gamma P^{\theta, c} f(r, v) \right\},$$

gdzie

$$\begin{aligned} P^{\theta,c} f(r, v) &= \mathbb{E} \left(f(R_{l+1}, V_{n+1}) | R_l = r, V_l = v, \theta_l = \theta, C_l = c \right) \\ &= \mathbb{E} \left(f \left(R_{l+1}, \theta_0 r_0 (v - c) + (v - c) \sum_{k=1}^{m} \theta_k R_{l+1,k} \right) | R_l = r \right) . \end{aligned}$$

Mamy $\mathcal{A}(0)(r, v) = v^{\alpha}$,

$$\begin{aligned} &\mathcal{A}^2(0)(r, v) \\ &= \sup_{0 \leq c \leq v} \left\{ c^{\alpha} + \gamma \sup_{\theta \in \varDelta_m} \mathbb{E} \left(\left(\theta_0 r_0 + \sum_{k=1}^{m} \theta_k R_{l+1,k} \right)^{\alpha} | R_l = r \right) (v - c)^{\alpha} \right\} . \end{aligned}$$

Niech

$$\psi_1(r) := \sup_{\theta \in \varDelta_m} \mathbb{E} \left(\left(\theta_0 r_0 + \sum_{k=1}^{m} \theta_k R_{l+1,k} \right)^{\alpha} | R_l = r \right)$$

oraz niech $\hat{\theta}^{(1)}(r) \in \varDelta_m$ będzie takie, że

$$\psi_1(r) = \mathbb{E}\left(\left(\hat\theta_0^{(1)} r_0 + \sum_{k=1}^m \hat\theta_k^{(1)} R_{l+1,k}\right)^\alpha |R_l = r\right).$$

Wtedy

$$\mathcal{A}^2(0)(r,v) = h\left(\gamma\psi_1(r)\right) v^\alpha$$

oraz

$$\mathcal{A}^3(0)(r,v) = \sup_{0\le c\le v}\left\{c^\alpha + \gamma\,\psi_2(r)(v-c)^\alpha\right\},$$

gdzie

$$\psi_2(r) := \sup_{\theta\in\varDelta_m} \mathbb{E}\left(\left(\theta_0 r_0 + \sum_{k=1}^m \theta_k R_{l+1,k}\right)^\alpha h\left(\gamma\psi_1(R_{l+1})\right)|R_l = r\right).$$

Stąd mamy następujący wynik.

Lemat 3.2 *Dla* $n = 0, 1, \ldots,$

$$\mathcal{A}^n(0)(r,v) = v^\alpha \psi_{n-1}(r), \qquad r\in D,\ v\ge 0,$$

gdzie $\psi_{-1}\equiv 0$ *oraz*

$$\psi_n(r) := \sup_{\theta\in\varDelta_m} \mathbb{E}\left(\left(\theta_0 r_0 + \sum_{k=1}^m \theta_k R_{l+1,k}\right)^\alpha h\left(\gamma\psi_{n-1}(R_{l+1})\right)|R_l = r\right),$$

a h *jest zdefiniowane wzorem* (3.3).

Zachodzi następujący wynik.

Twierdzenie 3.6 (*i*) *Granica* v_∞ *jest dana wzorem*

$$v_\infty(r,v) = v^\alpha \psi(r), \qquad r\in D,\ \ v\ge 0,$$

gdzie

$$\psi(r) = \lim_{n\to\infty}\psi_n(r), \qquad r\in D,$$

jest jedynym minimalnym nieujemnym rozwiązaniem równania

$$\varphi(r) := \sup_{\theta\in\varDelta_m} \mathbb{E}\left(\left(\theta_0 r_0 + \sum_{k=1}^m \theta_k R_{l+1,k}\right)^\alpha h\left(\gamma\varphi(R_{l+1})\right)|R_l = r\right). \qquad (3.12)$$

(*ii*) *Załóżmy, że* (R_n) *przyjmuje wartości w* $D = [0,M]^m$*, gdzie* $M < \infty$*. Jeżeli*

$$\gamma M^\alpha < 1, \qquad (3.13)$$

to minimalne rozwiązanie ψ *równania* (3.12) *jest funkcją ograniczoną na* D*. Ponadto optymalna strategia* $u_\infty = (c_\infty, \theta_\infty)$ *dana jest wzorem*

$$c_\infty(r,v) = g\left(\gamma\,\psi(r)\right)v = \left(1 + (\gamma\,\psi(r))^{1/(1-\alpha)}\right)^{-1} v, \quad r \in D,\, v \geq 0,$$

oraz

$$\psi(r) = \mathbb{E}\left(\left(\theta_{\infty,0} r_0 + \sum_{k=1}^{m} \theta_{\infty,k} R_{l+1,k}\right)^\alpha h\left(\gamma\psi(R_{l+1})\right)|R_l = r\right).$$

Dowód. Pierwsza część wynika z Lematu 3.2. Aby pokazać drugą część zauważmy, że

$$\left(\theta_0 r_0 + \sum_{k=1}^{m} \theta_k R_{l+1,k}\right)^\alpha \leq M^\alpha, \qquad \mathbb{P}-a.s.\ \forall\,\theta \in \Delta_m.$$

Stąd

$$\begin{aligned} 0 \leq \psi_n(r) &= M^\alpha \mathbb{E}\left(h(\gamma\,\psi_{n-1}(R_{l+1}))|R_l = r\right) \\ &\leq M^\alpha \left(1 + \gamma\,\|\psi_{n-1}\|_\infty\right) \end{aligned}$$

i możemy argumentować jak w dowodzie Twierdzenia 3.4.

3.7 Problem z ergodycznym funkcjonałem zysku

Równania Howarda–Bellmana dla modelu rozważnego w Rozdziale 3.3 są postaci

$$\begin{aligned} g(r,v) &= \sup_{0\leq c\leq v} \mathbb{E}\left(g(R_{n+1},V_{n+1})|R_n = r, V_n = v, C_n = c\right) \\ &= \sup_{0\leq c\leq v} Pg(\cdot, r(v-c))(r), \qquad r \in D,\ v \geq 0, \end{aligned}$$

oraz

$$g(r,v) + V(r,v) = \mathcal{A}V(r,v), \qquad r \in D,\ v \geq 0,$$

gdzie

$$\mathcal{A}V(r,v) = \sup_{0\leq c\leq v} \left\{c^\alpha + PV(\cdot, r(v-c))(r)\right\}.$$

Szukamy h i V postaci

$$g(r,v) = v^\alpha(r), \qquad V(r,v) = v^\alpha b(v).$$

Mamy

$$v^\alpha a(r) = \sup_{0\leq c\leq v} r^\alpha (v-c)^\alpha Pa(r) = r^\alpha v^\alpha Pa(r)$$

oraz

$$\begin{aligned} v^\alpha\left(a(r) + b(r)\right) &= \sup_{0\leq c\leq v} \left\{c^\alpha + r^\alpha (v-c)^\alpha P(a+b)(r)\right\} \\ &= v^\alpha h\left(r^\alpha P(a+b)(r)\right)). \end{aligned}$$

Tak więc

$$a(r) + b(r) = h(r^\alpha P(a+b)(r)), \qquad a(r) = r^\alpha Pa(r).$$

Niestety $a \equiv 0$ i $b(r) = h(r^\alpha Pb(r))$.

3.8 Ciągłe wersje modeli dyskretnych

Modele w czasie ciągłym otrzymujemy przechodząc do granicy $\Delta t \downarrow 0$. Właściwą wrersją w czasie dyskretnym jest

$$V_{t+\Delta t} = e^{r_t \Delta t} \left(V_t - C_t \Delta t \right). \tag{3.14}$$

Mamy

$$\frac{V_{t+\Delta t} - V_t}{\Delta t} = \frac{e^{r_t \Delta t} - 1}{\Delta t} V_t - C_t e^{r_t \Delta t}.$$

Przechodząc z $\Delta t \downarrow 0$ otrzymujemy

$$V_t' = r_t V_t - C_t. \tag{3.15}$$

Uwaga 3.1 Zauważmy, że model

$$V_t' = r_t (V_t - C_t)$$

nie ma poprawnej interpretacji ponieważ gdy $r_t \leq 0$ to wzrost C_t powoduje wzrost kapitału V_t!

W celu maksymalizacji funkcjonału

$$J(C, r_0, V_0) := \mathbb{E} \int_0^S e^{-\gamma t} C_t^\alpha dt, \tag{3.16}$$

gdzie S jest momentem bankructwa, powinniśmy sformułować funkcjonał zysku w czasie dyskretnym w następującej postaci

$$\mathbb{E} \sum_{t=0}^{\infty} \left(e^{-\gamma} \right)^t C_t^\alpha.$$

3.8.1 Formalne sformułowanie

Niech (r_t) będzie łańcuchem Markowa o wartościach w zbiorze ograniczonym $D \subset [0, M]$ z półgrupą przejścia Q_t. Dla zadanego $\varepsilon > 0$ niech $P^\varepsilon = P_\varepsilon$ oraz niech $r^\varepsilon = (r_n^\varepsilon;\ n = 0, 1, \ldots)$ będzie łańcuchem Markowa z operatorem przejścia Q_ε, to znaczy

$$Q_\varepsilon \psi(r) = \mathbb{E} \left(\psi(r_{n+1}^\varepsilon | r_n^\varepsilon = r \right).$$

Wówczas

$$R_n^\varepsilon := e^{\varepsilon r_n^\varepsilon}$$

jest łańcuchem Markowa z operatorem przejćia

$$\begin{aligned} P^\varepsilon \psi(R) &= \mathbb{E} \left(\psi \left(e^{\varepsilon r_{n+1}^\varepsilon} \right) | e^{\varepsilon r_n^\varepsilon} = R \right) \\ &= Q_\varepsilon \circ \tau_\varepsilon \psi \left(\frac{1}{\varepsilon} \log R \right), \end{aligned}$$

gdzie

$$\tau_\varepsilon \psi(x) = \psi\left(e^{\varepsilon x}\right).$$

Niech

$$V_{n+1}^\varepsilon = e^{r_n^\varepsilon \varepsilon}\left(V_n^\varepsilon - C_n\right) = R_n^\varepsilon\left(V_n^\varepsilon - C_n\right),$$

$$J_\varepsilon(C, R_0, V_0) = \mathbb{E}\sum_{n=0}^{\infty} e^{-\gamma\varepsilon n} C_n^\alpha,$$

oraz niech

$$W_\varepsilon(R, v) = \sup J_\varepsilon(C, R, v)$$

i

$$W(r, v) = \sup J(C, r, v),$$

gdzie supremum jest po wszystkich dopuszczalnych strategiach konsumpcji, są funkcjami wartości dla problemów z czasem dyskretnym i ciągłym. Oznaczmy przez $\widehat{C}^\varepsilon$ i $\widehat{C}$ optymalne konsumpcje (o ile istnieją) dla problemów w czasie dyskretnym i ciągłym. Tak więc

$$W_\varepsilon(r, v) = J_\varepsilon(\widehat{C}^\varepsilon, r, v)$$

i

$$W(r, v) = J(\widehat{C}, r, v).$$

Załóżmy, że $r = (r_t)$ jest procesem Markowa o wartościach w $E = (-\infty, \delta)$, gdzie $\delta < \gamma/\alpha$. Zauważmy, że r^ε również przyjmuje wartości w E, a więc $R_n^\varepsilon = e^{\varepsilon r_n^\varepsilon}$ przyjmuje wartości w $[0, M_\varepsilon]$ gdzie $M_\varepsilon = e^{\varepsilon\delta}$. Oczywiście

$$J_\varepsilon(R, v) = \mathbb{E}\sum_{n=0}^{\infty} \gamma_\varepsilon^n C_n^\alpha,$$

gdzie $\gamma_\varepsilon := e^{-\gamma\varepsilon}$. Zauważmy, że

$$\gamma_\varepsilon M_\varepsilon^\alpha = e^{\varepsilon(\delta\alpha - \gamma)} < 1.$$

Tat wię z Twierdzenia 3.4, W_ε jest skończone,

$$W_\varepsilon(R, v) = \psi_\varepsilon(R) v^\alpha,$$

gdzie ψ_ε jest minimalnym rozwiązaniem nieujemnym równania

$$\begin{aligned}\varphi(R) &= \left(1 + \left(\gamma_\varepsilon\, R^\alpha\, P^\varepsilon \varphi(R)\right)^{1/(1-\alpha)}\right)^{1-\alpha} \\ &= \left(1 + \left(e^{-\gamma\varepsilon}\, R^\alpha\, Q_\varepsilon \circ \tau_\varepsilon \varphi\left(\frac{1}{\varepsilon}\log R\right)\right)^{1/(1-\alpha)}\right)^{1-\alpha}, \quad R \in [0, M_\varepsilon].\end{aligned}$$

Używając podstawienia $R = e^{\varepsilon r}$ otrzymamy

$$\tilde{W}_\varepsilon(r,v) := W_\varepsilon(e^{\varepsilon r}, v) = v^\alpha \tilde{\psi}_\varepsilon(r),$$

gdzie $\tilde{\psi}_\varepsilon$ jest mninimalnym nieujemnym rozwiązaniem równania

$$\varphi(r) = \left(1 + \left(e^{\varepsilon(\alpha r - \gamma)} Q_\varepsilon \circ \tau_\varepsilon \varphi\left(r\right)\right)^{1/(1-\alpha)}\right)^{1-\alpha}, \qquad r \in (-\infty, \delta).$$

Niech

$$F_\varepsilon \varphi(r) := \left(1 + \left(e^{\varepsilon(\alpha r - \gamma)} Q_\varepsilon \circ \tau_\varepsilon \varphi\left(r\right)\right)^{1/(1-\alpha)}\right)^{1-\alpha}.$$

Wtedy $\tilde{\psi}_\varepsilon$ jest minimalnym rozwiązaniem

$$(F_\varepsilon - I)\varphi \equiv 0.$$

Zauważmy, że $F_\varepsilon \to F$, gdzie

$$F\varphi(r) = \left(1 + \left(\varphi\left(r\right)\right)^{1/(1-\alpha)}\right)^{1-\alpha}.$$

Oczywiście jedynym rozwiązaniem równania $F\varphi = \varphi$ jest $\varphi \equiv +\infty$.

Pokażemy, że

$$\lim_{\varepsilon \to 0} \varepsilon^{1-\alpha} \tilde{W}_\varepsilon(r,v) = W(r,v)$$

W tym celu policzymy granicę

$$\lim_{\varepsilon \to 0} \varepsilon^{1-\alpha} \tilde{\psi}_\varepsilon(r) =: \psi(r).$$

Mamy

$$\begin{aligned}\varepsilon^{1-\alpha} F_\varepsilon \varphi(r) &= \left(\varepsilon + \left(e^{\varepsilon(\alpha r - \gamma)} Q_\varepsilon \circ \tau_\varepsilon \varepsilon^{1-\alpha} \varphi\left(r\right)\right)^{1/(1-\alpha)}\right)^{1-\alpha} \\ &= \tilde{F}_\varepsilon \varepsilon^{1-\alpha} \varphi(r).\end{aligned}$$

Niech

$$\phi_\varepsilon(r) := \varepsilon^{1-\alpha} \tilde{\psi}_\varepsilon(r).$$

Wtedy

$$0 = \frac{\tilde{F}_\varepsilon \phi_\varepsilon - \phi_\varepsilon}{\varepsilon} = \frac{\tilde{F}_\varepsilon - \tilde{F}_0}{\varepsilon} \phi_\varepsilon.$$

Stąd

$$\begin{aligned}\psi(r) &= \frac{d}{d\varepsilon} \tilde{F}_\varepsilon \psi(r)|_{\varepsilon=0} \\ &= (1-\alpha)\left(\psi(r)\right)^{-\alpha/(1-\alpha)} \left(1 + \frac{1}{1-\alpha}\left(\psi(r)\right)^{\alpha/(1-\alpha)}\right)\end{aligned}$$

3.9 Kontrprzykład

Podamy przykład sterowanego łańcuchu Markowa oraz funkcjonału kosztu, takich że dla funkcji kosztu v_∞ zachodzi

$$v_\infty \neq \mathcal{A}\, v_\infty.$$

Niech

$$E = \{(j,k)\colon k = 1,2,\ldots, \quad j = 1,\ldots,k\}$$

oraz niech $U = \{1,2,\ldots\}$. Dla $x = (j,k) \in E$ oraz $u \in U$ połóżmy

$$\begin{aligned} q(x,u) &= 1 \quad \text{dla } j < k, \\ q(x,u) &= 0 \quad \text{dla } j = k. \end{aligned}$$

Ponadto niech

$$\begin{aligned} p^u((1,1),(1,u)) &= \frac{1}{u}, \qquad u \in U \setminus \{1\}, \\ p^u((1,1),(1,1)) &= 1 - \frac{1}{u}, \qquad u \in U, \\ p^u((j,k),(j+1,k)) &= 1, \qquad j \leq k-1,\ u \in U, \\ p^u((k,k),(k,k)) &= 1, \qquad k \geq 2,\ u \in U. \end{aligned}$$

Wówczas

$$\begin{aligned} &\mathcal{A}^n(0)((1,k)) = \min\{n, k-1\}\colon k \geq 2,\ n = 1,2,\ldots\} \\ &\mathcal{A}^n(0)((1,1)) = 0, \qquad n = 1,2,\ldots. \end{aligned}$$

Tak więc

$$v_\infty((1,k)) = k-1, \quad k \geq 2 \qquad \text{i} \qquad v_\infty((1,1)) = 0.$$

Z drugiej strony

$$\mathcal{A} v_\infty((1,1)) = \inf_{k\geq 2} \frac{k-1}{k} = \frac{1}{2} \neq 0 = v_\infty((1,1)).$$

Dla dowolnej strategii $\pi = (u_1, u_2, \ldots)$ mamy

$$\begin{aligned} J_\infty(\pi,(1,1)) &= \frac{1}{u_0}(u_0 - 1) + \left(1 - \frac{1}{u_0}\right)\frac{1}{u_1}(u_1 - 1) + \ldots \\ &= \left(1 - \frac{1}{u_0}\right) + \left(1 - \frac{1}{u_0}\right)\left(1 - \frac{1}{u_1}\right) + \ldots. \end{aligned}$$

Jest on minimalny dla $u_0 = u_1 = \ldots = 2$. Stąd dla $\widehat{\pi} = (2,2,\ldots)$ mamy

$$\inf_\pi J_\infty(\pi,(1,1)) = J_\infty(\widehat{\pi},(1,1)) = 1.$$

4

Zagadnienie liniowego regulatora

Zagadnienie liniowego regulatora nazywa się również *problemem liniowo-kwadratowym* ponieważ układ sterowany jest liniowy a funkcjonał kosztu jest kwadratowy. Tak więc dynamika procesu zdefiniowana jest rekurencyjnie

$$X_{n+1} = \Psi X_n + \Phi u_n + C\xi_n, \quad n = 0, 1, \dots,$$

gdzie X_0 zadana zmienna losowa, a Ψ, Φ i C są macierzami wymiarów odpowiednio $d\times d$, $d\times l$ i $d\times r$. Zakładamy, że zmienne losowe $\xi_0, \xi_1, \dots$ są wzajemnie niezależne, niezależne od X_0, mają średnią 0 i macierz kowariancji I. Ustalmy skończony horyzont N. Funkcjonał kosztu jest postaci

$$J_N(\pi, X_0) = \mathbb{E}\left\{\sum_{n=0}^{N-1} \gamma^n \left[\langle QX_n, X_n\rangle + \langle Ru_n, u_n\rangle\right] + \gamma^N \langle K_0 X_N, X_N\rangle\right\}, \tag{4.1}$$

gdzie Q, R, K_0 są symetrycznymi nieujemnie określonymi macierzami, a $\gamma > 0$ jest stałą dyskontującą. Będziemy zakładali, że R jest macierzą odwracalną. Przestrzenią stanów jest $\mathbb{R}^d$ a przestrzenią parametrów sterujących jest $\mathbb{R}^l$. Przez $\langle\cdot,\cdot\rangle$ oznaczamy iloczyn skalarny zarówno w $\mathbb{R}^d$ jak i w $\mathbb{R}^l$. Tak więc

$$\langle x, y\rangle = \sum_i x_i y_i.$$

Przestrzeń liniową macierzy o wymiarze $n \times m$ oznaczamy przez $M(n \times m)$. Przestrzeń macierzy symetrycznych nieujemnych wymiaru $n \times n$ oznaczamy przez $M_s^+(n)$.

4.1 Zasadnicze twierdzenie

Do sformułowania głównego twierdzenia podającego postać optymalnej strategii i wielkość minimalnego kosztu potrzebne mam będą dwa lematy.

Lemat 4.1 *Niech A i B będą operatorami liniowymi na przestrzeni liniowej E. Jeżeli operator $I+AB$ jest odwracalny, to operator $I+BA$ jest odwracalny. Ponadto*
$$(I+BA)^{-1} = I - B(I+AB)^{-1}A.$$

Dowód Niech
$$L = I - B(I+AB)^{-1}A.$$

Wtedy

$$\begin{aligned}(I+BA)L &= I + BA - (I+BA)B(I+AB)^{-1}A \\ &= I + BA - B(I+AB)(I+AB)^{-1}A \\ &= I + BA - BA = I.\end{aligned}$$

Podobnie

$$\begin{aligned}L(I+BA) &= I + BA - B(I+AB)^{-1}A(I+BA) \\ &= I + BA - B(I+AB)^{-1}(I+AB)A \\ &= I + BA - BA = I.\end{aligned}$$

□

Lemat 4.2 *Jeżeli $A, B \in M_s^+(n)$, to macierz $I+AB$ jest odwracalna i $B(I+AB)^{-1} \in M_s^+(n)$.*

Dowód Ponieważ $A \in M_s^+(n)$ to $A = \sqrt{A}\sqrt{A}$, gdzie $\sqrt{A} \in M_s^+(n)$. Następnie, macierz $I + \sqrt{A}B\sqrt{A}$ jest symetryczna i ściśle dodatnio określona. Jest więc odwracalna, a więc na podstawie poprzedniego lematu odwracalna jest macierz
$$I + AB = I + \sqrt{A}(\sqrt{A}B).$$

Co więcej

$$(I+AB)^{-1} = (I + \sqrt{A}\sqrt{A}B)^{-1} = I - \sqrt{A}(I + \sqrt{A}B\sqrt{A})^{-1}\sqrt{A}B.$$

Stąd
$$B(I+AB)^{-1} = B - B\sqrt{A}(I + \sqrt{A}B\sqrt{A})^{-1}\sqrt{A}B,$$
czyli macierz $B(I+AB)^{-1}$ jest symetryczna. Niech $x \in \mathbb{R}^d$ oraz

$$y = (I+AB)^{-1}x.$$

Wówczas mamy

$$\begin{aligned}\langle B(I+AB)^{-1}x, x\rangle &= \langle By, (I+AB)y\rangle \\ &= \langle By, y\rangle + \langle By, \sqrt{A}\sqrt{A}By\rangle \\ &= |\sqrt{B}y|^2 + |\sqrt{A}By|^2 \geq 0,\end{aligned}$$

co kończy dowód lematu. □.

Stosując ostatni lemat do $B = K$ i $A = \gamma\Phi R^{-1}\Phi^*$ wnioskujemy, że wzór

$$\mathcal{A}_\gamma(K) = Q + \gamma\Psi^* K(I + \gamma\Phi R^{-1}\Phi^* K)^{-1}\Psi, \qquad K \in M_s^+(d)$$

definiuje odwzorowanie z $M_s^+(d)$ w $M_s^+(d)$. Iteracje odwzorowania $\mathcal{A}_\gamma$ oznaczać będziemy przez $\mathcal{A}_\gamma^n$, $n = 0, 1, \ldots$.

Jeżeli $K \in M_s^+(d)$, to $R + \Phi^* K\Phi \in M_s^+(l)$ i $(R + \Phi^* K\Phi)^{-1} \in M_s^+(l)$. Ponadto

$$\mathcal{B}_\gamma(K) = -\gamma(R + \gamma\Phi^* K\Phi)^{-1}\Phi^* K\Psi \in M(d \times l).$$

Twierdzenie 4.1 *Niech $K_n = \mathcal{A}_\gamma^n(K_0)$, $n = 0, 1, \ldots$. Wtedy minimalny koszt w zagadnieniu liniowego regulatora wynosi*

$$\mathbb{E}\,\langle K_N X_0, X_0\rangle + \sum_{n=0}^{N-1} \gamma^n \operatorname{Trace} C^* K_n C,$$

a strategia

$$\widehat{\pi} = (\mathcal{B}_\gamma(K_{N-1})x_0, \ldots, \mathcal{B}_\gamma(K_0)x_{N-1})$$

jest optymalna.

Uwaga Zauważmy, że optymalna strategia nie zależy od szumu, w tym sensie, że macierze $\mathcal{B}_\gamma(K_n)$, $n = N-1, \ldots, 0$, nie zależą od macierzy C. W szczególności dla układu deterministycznego, to jest gdy $C = 0$, strategia optymalna jest identyczna ze strategią dla układu z szumem. Fakt ten, w teorii sterowania, nosi nazwę *zasady niezmienniczości* (*invariance principle*).

Dowód twierdzenia poprzedzimy dwoma lematami.

Lemat 4.3 *Jeżeli $R \in M_s^+(l)$ jest macierzą odwracalną to*

$$\langle Ru, u\rangle + \langle a, u\rangle \geq -\frac{1}{4}\langle R^{-1}a, a\rangle, \qquad \forall\, u, a \in \mathbb{R}^l.$$

Ponadto równość zachodzi wtedy i tylko wtedy gdy

$$u = \widehat{u} := -\frac{1}{2}R^{-1}a.$$

Dowód Dla dowolnego u mamy

$$\langle R(u - \widehat{u}), (u - \widehat{u})\rangle \geq 0$$

i równość zachodzi wtedy i tylko wtedy gdy $u = \widehat{u}$. Ponieważ

$$\langle R(u - \widehat{u}), u - \widehat{u}\rangle = \langle Ru, u\rangle + \langle a, u\rangle + \frac{1}{4}\langle R^{-1}a, a\rangle$$

dowód lematu jest zakończony. □

Lemat 4.4 *Jeżeli ξ jest wektorem losowym w $\mathbb{R}^n$ o wartości oczekiwanej 0 i macierzy kowariancji S, to dla dowolnej macierzy $K \in M(n \times n)$ mamy*

$$\mathbb{E}\,\langle K\xi, \xi\rangle = \operatorname{Trace} KS.$$

Dowód Niech $e_1, \dots, e_n$ będzie bazą ortonormalną przestrzeni $\mathbb{R}^n$. Wtedy

$$\langle K\xi, \xi\rangle = \sum_{j=1}^{n} \langle K\xi, e_j\rangle\langle \xi, e_j\rangle.$$

Stąd

$$\begin{aligned}
\mathbb{E}\,\langle K\xi, \xi\rangle &= \mathbb{E}\sum_{j=1}^{n} \langle K\xi, e_j\rangle\langle \xi, e_j\rangle \\
&= \sum_{j=1}^{n} \langle Se_j, K^* e_j\rangle \\
&= \sum_{j=1}^{n} \langle KSe_j, e_j\rangle = \operatorname{Trace} KS.
\end{aligned}$$

□

Dowód Twierdzenia 4.1 Skorzystamy z twierdzenia ogólnego Bellmana dla sterowania w czasie dyskretnym, to jest z Twierdzenia 2.2. W twierdzeniu tym dla $x \in \mathbb{R}^q$, $u \in \mathbb{R}^l$ oraz $\xi \in \mathbb{R}^r$ przyjmiemy

$$\begin{aligned}
F(x,u,\xi) &= \Psi x + \Phi u + C\xi \\
q(x,u) &= \langle Qx, x\rangle + \langle Ru, u\rangle, \\
r_N(x) &= \langle K_0 x, x\rangle.
\end{aligned}$$

Wystarczy udowodnić, że dla dowolnych $x \in \mathbb{R}^d$ i $K \in M_s^+(d)$ mamy

$$\begin{aligned}
&\inf_{u\in\mathbb{R}^l} \{q(x,u) + \gamma\,\mathbb{E}\,\langle KF(x,u,\xi_0), F(x,u,\xi)\rangle\} \\
&= \inf_{u\in\mathbb{R}^l} \{\langle Qx, x\rangle + \langle Ru, u\rangle \\
&\qquad + \gamma\mathbb{E}\,\langle K\Psi x + K\Phi u + KC\xi_0, \Psi x + \Phi u + C\xi_0\rangle\} \\
&= \langle \mathcal{A}_\gamma(K)x, x\rangle + \gamma\operatorname{Trace} C^* KC,
\end{aligned}$$

gdzie infinum jest osiągalne dla $u = \mathcal{B}_\gamma(K)x$. Z Lematu 4.4 i z faktu, że ξ_j mają zerową wartość oczekiwaną i kowariancje I mamy

$$\begin{aligned}
&\mathbb{E}\,\langle K\Psi x + K\Phi u + KC\xi_0, \Psi x + \Phi u + C\xi_0\rangle \\
&= \langle K\Psi x + K\Phi u, \Psi x + \Phi u\rangle + \operatorname{Trace} C^* KC \\
&= \langle K\Psi x, \Psi x\rangle + \operatorname{Trace} C^* KC + 2\langle K\Psi x, \Phi u\rangle + \langle K\Phi u, \Phi u\rangle.
\end{aligned}$$

Z Lematu 4.3 infimum po u wyrażenia

$$\langle Ru, u\rangle + \gamma\langle K\Phi u, \Phi u\rangle + 2\gamma\langle K\Psi x, \Phi u\rangle = \langle (R + \gamma\Phi^* K\Phi)u, u\rangle + \langle 2\gamma\Phi^* K\Psi x, u\rangle$$

jest osiągalne w punkcie

$$\widehat{u} = -\frac{1}{2}\left(R + \gamma\Phi^* K \Phi\right)^{-1} 2\gamma\Phi^* K \Psi x = \mathcal{B}_\gamma(K)x$$

i wynosi

$$\begin{aligned} &-\frac{1}{4}\langle (R + \gamma\Phi^* K\Phi)^{-1} 2\gamma\Phi^* K\Psi x, 2\gamma\Phi^* K\Psi x\rangle \\ &= -\gamma^2\langle (R + \gamma\Phi^* K\Phi)^{-1}\Phi^* K\Psi x, \Phi^* K\Psi x\rangle. \end{aligned}$$

Stąd

$$\begin{aligned} &\inf_{u\in\mathbb{R}^l} \{q(x,u) + \gamma\mathbb{E}\,\langle KF(x,u,\xi_0), F(x,u,\xi)\rangle\} \\ &= \langle (Q + \gamma\Psi^* K\Psi)x, x\rangle - \gamma^2\langle (R + \gamma\Phi^* K\Phi)^{-1}\Phi^* K\Psi x, \Phi^* K\Psi x\rangle \\ &\qquad\qquad +\gamma \mathrm{Trace}\, C^* K C \\ &= \left\langle \left(Q + \gamma\Psi^* K\Psi - \gamma^2\Psi^* K\Phi(R + \gamma\Phi^* K\Phi)^{-1}\Phi^* K\Psi\right) x, x\right\rangle \\ &\qquad +\gamma \mathrm{Trace}\, C^* K C. \end{aligned}$$

Należy więc pokazać, że

$$\begin{aligned} &Q + \gamma\Psi^* K\Psi - \gamma^2\Psi^* K\Phi(R + \gamma\Phi^* K\Phi)^{-1}\Phi^* K\Psi \\ &= \mathcal{A}_\gamma(K) =: Q + \gamma\Psi^* K(I + \gamma\Phi R^{-1}\Phi^* K)^{-1}\Psi. \end{aligned}$$

Mamy

$$\begin{aligned} &\gamma\Psi^* K\Psi - \gamma^2\Psi^* K\Phi(R + \gamma\Phi^* K\Phi)^{-1}\Phi^* K\Psi \\ &= \Psi^* K\left(\gamma I - \gamma^2\Phi(R + \gamma\Phi^* K\Phi)^{-1}\Phi^* K\right)\Psi. \end{aligned}$$

Stosując Lematy 4.1 i 4.2 dla

$$A = \gamma R^{-1}\Phi^* K, \qquad B = \Phi$$

otrzymujemy

$$\begin{aligned} \gamma I - \gamma^2\Phi(R + \gamma\Phi^* K\Phi)^{-1}\Phi^* K &= \gamma\left\{I - \gamma\Phi(R + \gamma\Phi^* K\Phi)^{-1}\Phi^* K\right\} \\ &= \gamma\left\{I - \Phi(I + \gamma R^{-1}\Phi^* K\Phi)^{-1}\gamma R^{-1}\Phi^* K\right\} \\ &= \gamma\left\{I - B\left(I + AB\right)^{-1} A\right\} \\ &= \gamma\left(I + BA\right)^{-1} \\ &= \gamma\left(I + \gamma\Phi R^{-1}\Phi^* K\right)^{-1}, \end{aligned}$$

co daje żądaną równość. □

4.2 Sterowanie i stan zależne od szumu

Wyniki dotyczące problemu liniowego regulatora mogą być uogólnione na tak zwany przypadek *liniowego regulatora ze sterowaniem i stanem zależnymi od szumu*, to znaczy gdy dynamika procesu jest następująca

$$X_{n+1} = \Psi X_n + A(X_n, \xi_n^1) + \Phi u_n + B(u_n, \xi_n^2) + C\xi_n,$$

gdzie A i B są odwzorowaniami dwuliniowymi z $\mathbb{R}^d \times \mathbb{R}^{r_1}$ w $\mathbb{R}^d$ i z $\mathbb{R}^l \times \mathbb{R}^{r_2}$ w $\mathbb{R}^d$. Zakładamy, że zmienne losowe ξ_n^1, ξ_m^2, ξ_r, $n, m, r = 0, 1, \ldots$, o wartościach w $\mathbb{R}^{r_1}$, $\mathbb{R}^{r_2}$ i $\mathbb{R}^r$, są niezależne, o zerowych średnich i kowariancjach odpowiednio I_{r_1}, I_{r_2} i I_r. Funkcjonał kosztu dany jest wzorem (4.1).

Ważnym przypadkiem jest system z rzeczywistymi (ξ_n^1) i (ξ_n^2). Wtedy

$$A(x_n, \xi_n^1) = \xi_n^1 \tilde{A} x_n \quad \text{i} \quad B(u_n, \xi_n^1) = \xi_n^2 \tilde{B} u_n,$$

z odpowiednio dobranymi macierzami $\tilde{A}$ i $\tilde{B}$.

Niech $M_s(n)$ oznacza przestrzeń macierzy symetrycznych wymiaru $n \times n$. Przypomnijmy, że $M_s^+(n)$ to przestrzeń macierzy symetrycznych nieujemnych.

Niech $\mathcal{G}_1$ i $\mathcal{G}_2$ będą odwzorowaniami liniowymi z $M_s(d)$ w $M_s(d)$ i z $M_s(l)$ w $M_s(l)$ danymi wzorami

$$\begin{aligned}
\langle \mathcal{G}_1(L)x, y\rangle &= \mathbb{E}\,\langle LA(x, \xi_0^1), A(y, \xi_0^1)\rangle, \qquad L \in M_s(d),\ x, y \in \mathbb{R}^d \\
\langle \mathcal{G}_2(L)x, y\rangle &= \mathbb{E}\,\langle LB(x, \xi_0^2), B(y, \xi_0^2)\rangle, \qquad L \in M_s(l),\ x, y \in \mathbb{R}^l.
\end{aligned}$$

Zauważmy, że $\mathcal{G}_1 \colon M_s^+(d) \to M_s^+(d)$ i $\mathcal{G}_2 \colon M_s(l) \to M_s(l)$.

Niech

$$\begin{aligned}
\mathcal{A}_1(K) &= Q + \mathcal{G}(K) + \Psi^* K \left(I + \Phi(R + \mathcal{G}_2(K))^{-1}\Phi^* K\right)^{-1} \Psi \\
\mathcal{B}_1(K) &= -\left(R + \Phi^* K \Phi + \mathcal{G}_2(K)\right)^{-1} \Phi^* K \Psi.
\end{aligned}$$

Zachodzi następujący analogon Twierdzenia 4.1

Twierdzenie 4.2 *Niech $K_n = \mathcal{A}_1^n(K_0)$. Wówczas minimalny koszt na przedziale czasowym $[0, N]$ wynosi*

$$\langle K_N x, x\rangle + \sum_{n=0}^{N-1} \operatorname{Trace} C^* K_n C,$$

a optymalna strategia dana jest wzorem

$$\widehat{u}_0(x) = \mathcal{B}_1(K_{N-1})x, \ldots, \widehat{u}_{N-1}(x) = \mathcal{B}_1(K_0)x, \qquad x \in \mathbb{R}^d.$$

Dowód twierdzenia jest bardzo podobny do dowodu Twierdzenia 4.1. Kluczowy Lemat 4.3 ma następującą postać.

Lemat 4.5 *Niech*

$$\begin{aligned}
F(x, u, \xi_1, \xi_2, \xi_3) &= \Psi x + A(x, \xi_1) + \Phi u + B(x, \xi_2) + C\xi_3 \\
q(x, u) &= \langle Qx, x\rangle + \langle Ru, u\rangle.
\end{aligned}$$

Wówczas dla dowolnych $K \in M_s^+(d)$ i $x \in \mathbb{R}^d$,

$$\begin{aligned}
&\inf_{u \in \mathbb{R}^l} \left\{ q(x, u) + \mathbb{E}\,\langle KF(x, u, \xi_0^1, \xi_0^2, \xi_0), F(x, u, \xi_0^1, \xi_0^2, \xi_0)\rangle \right\} \\
&= \langle \mathcal{A}_1(K)x, x\rangle + \operatorname{Trace} C^* K C.
\end{aligned}$$

Ponadto infimum jest osiągalne w punkcie $\widehat{u}(x) = \mathcal{B}_1(K)x$.

5

Stopowanie - horyzont skończony

W teorii optymalnego stopowania zadany proces Markowa $(X_n, \mathcal{X}_n)$ na przestrzeni stanów $(E, \mathcal{E})$. Celem jest znalezienie momentu zatrzymania τ, który maksymalizuje (minimalizuje) zadany funkcjonał J zysku (kosztu). Będziemy zakładali, że J dany jest wzorem

$$J(\tau, X_0) = \mathbb{E}\left(\sum_{n=0}^{\tau-1} \gamma^n q(X_n) + \gamma^\tau r(X_\tau)\right),$$

gdzie $q, r\colon E \to [0, +\infty)$ są mierzalnymi funkcjami nieujemnymi, a γ liczbą nieujemną. Od τ wymagamy by był on momentem Markowa względem filtracji $(\mathcal{X}_n)$. Rozważamy problem optymalnego stopowania ze skończonym horyzontem czasowym $N < +\infty$. Tak więc dodatkowo będziemy zakładali, że $\mathbb{P}(\tau \le N) = 1$. Klasę tych momentów Markowa będziemy oznaczali przez Σ_N.

Na początku założymy, że proces zadany jest rekurencyjnie, to znaczy, że

$$X_{n+1} = F(X_n, \xi_{n+1}), \qquad n = 0, 1, \dots$$

Powyżej, $\{\xi_n\}$ jest ciągiem niezależnych zmiennych losowych o tym samym rozkładzie w przestrzeni mierzalnej $(S, \mathcal{S})$, $\{X_n\}$ przyjmują wartości w mierzalnej przestrzeni stanów $(E, \mathcal{E})$, a F jest mierzalnym odwzorowaniem produktu $E \times S$ w E. Filtracja zadana jest proces (X_n), to znaczy

$$\mathcal{X}_n = \sigma(X_0, \dots, X_n), \qquad n = 0, 1, \dots.$$

Tak więc mament Markowa $\tau \in \Sigma_n$ spełnia warunek

$$\{\tau \le n\} \in \sigma(X_0, \dots, X_n), \qquad n = 0, 1, \dots.$$

5.1 Zasadnicze twierdzenia

Niech P będzie operatorem przejścia dla procesu $\{X_n\}$, to jest niech

$$Ph(x) := \mathbb{E}\, h(F(x,\xi_0)).$$

Zdefiniujemy operator Q formułą

$$Qh(x) = \max\left(q(x) + \gamma Ph(x), r(x)\right), \qquad x \in E,$$

gdy J jest zyskiem, a

$$Qh(x) = \min\left(q(x) + \gamma Ph(x), r(x)\right), \qquad x \in E,$$

gdy J jest kosztem.

Twierdzenie 5.1 *Jeżeli J jest zyskiem to moment zatrzymania*

$$\begin{aligned}\widehat{\tau} &:= \inf\{n \le N \colon Q^{N-n} r(X_n) \le r(X_n)\}\\ &= \inf\{n \le N \colon q(X_n) + \gamma P Q^{N-(n+1)} r(X_n) \le r(X_n)\}\end{aligned}$$

jest optymalny, to jest

$$J(\widehat{\tau}, X_0) = \sup_{\tau \in \Sigma_N} J(\tau, X_0).$$

Moment zatrzymania

$$\begin{aligned}\widehat{\tau} &:= \inf\{n \le N \colon Q^{N-n} r(X_n) \ge r(X_n)\}\\ &= \inf\{n \le N \colon q(X_n) + \gamma P Q^{N-(n+1)} r(X_n) \ge r(X_n)\}\end{aligned}$$

jest optymalny dla funkcjonału kosztu, czyli

$$J(\widehat{\tau}, X_0) = \inf_{\tau \in \Sigma_N} J(\tau, X_0).$$

Ponadto, dla funkcjonałów zysku i kosztu

$$J(\widehat{\tau}, X_0) = \mathbb{E}\, Q^N r(X_0).$$

Dowód Dowód polega na przeformułowanie problemu optymalnego stopowania na problem optymalnego sterowania. W tym celu zdefiniujemy nową przestrzeń stanów $\tilde{E} = E \cup \{\delta\}$, gdzie δ jest sztucznym punktem nie należącym do E. Niech $U = \{0, 1\}$ będzie dwuelementową przestrzenią sterowań. Nowy system $\{\tilde{X}_n\}$ zadany będzie przez następującą rodzine prawdopodobieństw przejścia $\{P^u;\ u \in U\}$,

$$P^0(x, \Gamma) = P(x, \Gamma), \qquad x \in E,\ \Gamma \in \mathcal{E},$$

$$P^0(x, \{\delta\}) = 0, \qquad P^1(x, \{\delta\}) = 1,$$

$$P^0(\delta, \{\delta\}) = P^1(\delta, \{\delta\}) = 1, \quad x \in E.$$

Określmy również nowy funkcjonał zysku

$$\tilde{J}(\pi, \tilde{X}_0) = \mathbb{E}\left(\sum_{n=0}^{N-1} \gamma^n \tilde{q}(\tilde{X}_n, u_n) + \gamma^N \tilde{r}(\tilde{X}_N)\right),$$

gdzie

$$\tilde{q}(x,0) = q(x), \quad q(x,1) = r(x), \quad \tilde{r}(x) = r(x), \qquad x \in E,$$
$$q(\delta, u) = 0, \quad r(\delta) = 0.$$

Ogólnie funkcje $h\colon E \to \mathbb{R}$ przedłużamy na $\tilde{E}$ kładąc $h(\delta) = 0$.

Przyjmijmy konwencje $\min \emptyset = N$. Zauważmy, że dwie strategie

$$\pi^1 = (u_1^1, \dots, u_{N-1}^1) \quad \text{oraz} \quad \pi^2 = (u_1^2, \dots, u_{N-1}^2),$$

takie że

$$\forall x \in \tilde{E}, \qquad \inf\left\{n \le N\colon u_n^1(x) = 1\right\} = \inf\left\{n \le N\colon u_n^2(x) = 1\right\}$$

dają ten sam zysk (koszt). Wynika to z tego, że

$$q(\delta, u) = 0 = r(\delta), \quad u \in U \qquad \text{i} \qquad P^1(x, \{\delta\}) = 1, \quad x \in \tilde{E}.$$

Dwie takie strategie będziemy utożsamiać, pisząc $\pi^1 \cong \pi^2$. Oczywiście $\cong$ jest relacją równoważności na zbiorze Π wszystkich strategii dla problemu $(\tilde{X}), \tilde{J}))$.

Zauważmy, że istnieje wzajemnie jednoznaczne odwzorowanie (bijekcja) j działające ze zbioru $\Pi/\cong$ w zbiór Σ_N, takie że

$$\tilde{J}(\pi, X_0) = J(j\pi, X_0).$$

Istotnie odwzorowanie

$$j\pi = \min\{n \le N-1\colon u_n = 1\}, \quad \pi = (u_0, \dots, u_{N-1})$$

ma żądane własności. Zauważmy, że

$$j^{-1}\tau = (u_n)\colon u_n = 0 \text{ dla } n < \tau \text{ i } u_n = 1 \text{ dla } n \in \{\tau, \dots, N-1\}.$$

Niech $\{\tilde{V}_n\}$ będą funkcjami wartości występującymi w Twierdzeniu (Bellmana) 2.2, oraz niech $\mathcal{A}$ będzie operatorem Bellmana dla problemu sterowania. Wówczas

$$\mathcal{A}h(x) = Qh(x), \qquad x \in \tilde{E}$$

Istotnie, gdy J jest zyskiem, to dla $x \in E$,

$$\begin{aligned}
\mathcal{A}h(x) &= \max_{u=0,1}\left(q(x,u) + \gamma P^u h(x)\right) \\
&= \max\left(q(x) + \gamma P h(x), r(x) + h(\delta)\right) \\
&= \max\left(q(x) + \gamma P h(x), r(x)\right) \\
&= Qh(x).
\end{aligned}$$

Zauważmy, że

$$\begin{aligned}\mathcal{A}h(\delta) &= \max\big(\tilde{q}(\delta,0)+\gamma Ph(\delta), \tilde{q}(\delta,1)+\gamma P^1 h(\delta)\big)\\ &= h(\delta) = 0 = Qh(\delta).\end{aligned}$$

Czyli mamy żądaną równość $\mathcal{A} = Q$.

Oczywiście to samo rozumowanie można zastosować dla funkcjonału kosztu. Stąd, dla problemu $((\tilde{X}_n), (\tilde{J}))$, funkcje $\{V_n\}$ spełniają, patrz Twierdzenie 2.2,

$$V_n = \gamma^n Q^{N-n} r.$$

Ponadto strategia $\widehat{\pi} = (\widehat{u}_0, \ldots, \widehat{u}_{N-1})$, która spełnia

$$Q^{N-n}r(x) = q(x, \widehat{u}_n(x)) + \gamma P^{\widehat{u}_n(x)} Q^{N-(n+1)} r(x), \qquad n = 0, \ldots, N-1$$

jest optymalna. Oczywiście

$$Q^{N-n}r(\delta) = 0 = q(\delta, u) + \gamma P^u Q^{N-(n+1)} r(\delta), \quad u = 0, 1.$$

Stąd wartość $\widehat{u}_n(\delta)$ może być dowolna. My przyjmujemy $\widehat{u}_n(\delta) = 1$. Ponieważ, dla $x \in E$,

$$q(x, 0) + \gamma P^0 Q^{N-(n+1)} r(x) = q(x) + \gamma P Q^{N-(n+1)} r(x),$$

a

$$q(x, 1) + \gamma P^1 \widehat{Q}^{N-(n+1)} r(x) = r(x),$$

mamy

$$\widehat{u}_n(x) = \begin{cases} 0, & \text{gdy } q(x) + \gamma P Q^{N-(n+1)} r(x) > r(x) \\ 1, & \text{gdy } q(x) + \gamma P Q^{N-(n+1)} r(x) \leq r(x). \end{cases}$$

Wystarczy więc pokazać, że strategia $(\tilde{u}_n) := j^{-1}\widehat{\tau}$ spełnia powyższy warunek. Wynika to z definicji $\widehat{\tau}$. Istotnie

$$\tilde{u}_n(X_n) = 0 \quad \leftrightarrow \quad n < \widehat{\tau} \quad \leftrightarrow \quad Q^{N-n} r(X_m) > r(X_m) \ \forall m \leq n,$$

a na to aby $Q^{N-n}r(x) > r(x)$ potrzeba i wystarcza by

$$q(x) + \gamma P Q^{N-(n+1)} r(x) > r(x).$$

□

5.2 Interpretacja

Twierdzenia o optymalnym stopowaniu mają bardzo prostą interpretację. Rozważmy przypadek funkcjonału zysku. Maksymalny średni zysk gdy mamy do dyspozycji horyzont czasowy N wynosi $Q^N r(X_0)$. Tutaj za X_0 wstawiamy zaobserwowaną wartość zmiennej losowej X_0 i podejmujemy decyzję; realizować zysk

$$J(0, X_0) = r(X_0)$$

lub odłożyć decyzję na później. Oczywiście odłożenie decyzji jest umotywowane tylko wtedy gdy maksymalny zysk $Q^N r(X_0)$ jest silnie większy od tego cobyśmy otrzymali natychmiast, czyli od $r(X_0)$. Jeżeli $Q^N r(X_0) > r(X_0)$ to odkładamy decyzję do czasu $t = 1$. Wówczas, mamy wartość X_1, horyzont $N - 1$ oraz maksymalny zysk

$$\gamma Q^{N-1} r(X_1) + q(X_0).$$

To co byśmy otrzymali zatrzymując się w momencie $t = 1$ wynosi

$$\gamma r(X_1) + q(X_0)$$

Powinniśmy więc rozumować jak dla czasu $t = 0$. Ogólnie stopujemy w chwili gdy maksymalny zysk w pozostałym przedziale czasowym $N - n$, to jest

$$\sum_{k=0}^{n-1} \gamma^k q(X_k) + \gamma^{N-n} Q^{N-n} r(X_n)$$

nie przekracza tego co byśmy dostali stopując w momencie n, to jest

$$\sum_{k=0}^{n-1} \gamma^k q(X_k) + \gamma^{N-n} r(X_n).$$

5.3 Stopowanie ciągu

Problem optymalnego stopowania na skończonym przedziale czasowym zilustrujemy następującym przykładem. Niech $\{X_0\}$ będzie ciągiem niezależnych zmiennych losowych o rozkładzie jednostajnym na odcinku $[0, 1]$. Ustalmy skończony horyzont czasowy N. Naszym celem jest znalezienie momentu Markowa $\widehat{\tau}$, który maksymalizuje funkcjonał

$$J(\tau, X_0) = \mathbb{E}\, \gamma^\tau X_\tau, \qquad \tau \in \Sigma_N,$$

gdzie $\gamma \in (0, +\infty)$.

W tym celu zdefiniujmy rekurencyjnie ciąg liczb

$$m_0 = 0, \qquad m_{n+1} = \frac{\gamma}{2}\left(1 + m_n^2\right), \quad n = 0, 1, \ldots.$$

Mamy następujące twierdzenie.

Twierdzenie 5.2 *Optymalny moment zatrzymania $\widehat{\tau}$ dany jest wzorem*

$$\widehat{\tau} = \min\{n \le N \colon X_n \ge m_{N-n}\}.$$

Ponadto maksymalny zysk spełnia

$$\max_{\tau \in \Sigma_N} \mathbb{E}\, X_\tau = \mathbb{E}\, X_{\widehat{\tau}} = m_{N+1}.$$

Dowód Dla powyższego problemu przestrzeń stanów $E = [0,1] = S$,

$$F(X_n, \xi_n) = \xi_n, \qquad n = 0, 1, \ldots.$$

Operator przejścia dany jest więc wzorem

$$Ph(x) = \mathbb{E}\, h(\xi_0) = \int_0^1 h(x)dx.$$

Dla funkcjonał u zysku mamy $q(x) = 0$, $r(x) = x$, $x \in [0,1]$, $\gamma = 1$. Tak więc operator Q dany jest wzorem

$$Qh(x) = \max\,\{\gamma Ph(x), x\} = \max\left\{\gamma \int_0^1 h(x)dx, x\right\}.$$

Wystarczy więc pokazać, że

$$Q^n r(x) = \max\{m_n, x\}, \qquad x \in [0,1],\ n = 0, 1, \ldots.$$

Wykażemy to stosując indukcje względem n. Ponieważ $m_0 = 0$,

$$Q^0 r(x) = r(x) = x = \max\{0, x\}.$$

Następnie, z założenia indukcyjnego,

$$\begin{aligned}
Q^{n+1} r(x) &= \max\left\{\gamma \int_0^1 Q^n r(x)dx, x\right\} \\
&= \max\left\{\gamma m_n^2 + \gamma \int_{m_n}^1 xdx, x\right\} \\
&= \max\left\{\gamma m_n^2 + \frac{\gamma}{2}\left(1 - m_n^2\right), x\right\} \\
&= \max\left\{\frac{\gamma}{2}\left(1 + m_n^2\right), x\right\} \\
&= \max\{m_{n+1}, x\}.
\end{aligned}$$

□

Oczywiście analogiczny problem można postawić dla funkcjonału kosztu. To znaczy, że teraz jesteśmy zainteresowani znalezieniem $\overline{\tau} \in \Sigma_N$, który minimalizuje $\mathbb{E}\,\gamma^\tau X_\tau$. Teraz

$$Qh(x) = \min \left\{ \gamma \int_0^1 h(x)dx, x \right\}, \qquad \in [0,1].$$

Stąd łatwo pokazać, że

$$Q^n r(x) = \min \{l_n, x\}, \qquad n = 0, 1, \ldots, \quad x \in [0,1],$$

gdzie ciąg $\{l_n\}$ dany jest rekurencyjnie

$$l_0 = 1, \qquad l_{n+1} = l_n - \frac{l_n^2}{2}, \quad n = 0, 1, \ldots.$$

mamy więc następujący rezultat

Twierdzenie 5.3 *Moment zatrzymania*

$$\overline{\tau} = \min \{n = 0, \ldots, N \colon X_n \le l_{N-n}\}$$

minimalizuje funkcjonał

$$\mathbb{E}\, X_\tau, \qquad \tau \in \Sigma_N.$$

Ponadto

$$\min_{\tau \in \Sigma_N} \mathbb{E}\, X_\tau = \mathbb{E}\, X_{\overline{\tau}} = l_{N+1}.$$

5.4 Strategia na egzamin

Naszym drugim przykładem zagadnienia optymalnego stopowania jest problem znalezienie najlepszej strategii na egzamin. Wyobraźmy sobie studenta, który zna odpowiedzi na K pytań spośród $W = K + M$. W czasie egzaminu, każdy student losuje (bez zwracania) jedno pytanie. Student zna pytania, które zostały wylosowane i ma prawo wyboru momentu podejścia do egzaminu. Jaka jest dla niego najlepsza strategia? A może wszystkie dają to samo prawdopodobieństwo zdania?

Niech (K_n, M_n) oznaczają odpowiednio liczbe pytań na które student potrafi odpowiedzieć, oraz liczbę pytań, na które nie potrafi odpowiedzieć po n-tym losowaniu. Oczywiście $K_0 = K$, $M_0 = M$. Ciąg $X_n = (K_n, M_n)$ jest łańcuchem Markowa na przestrzeni stanów

$$E = \{(k,m) \colon k = 0, 1, \ldots, K, \qquad m = 0, 1, \ldots, M\}.$$

Jego prawdopodobieństwa przejścia wynoszą

$$P\left((k,m), \{(k-1,m)\}\right) = \frac{k}{k+m}, \qquad k \ge 1,$$

$$P\left((k,m), \{(k,m-1)\}\right) = \frac{m}{k+m}, \qquad m \ge 1,$$

$$P\left((0,0),\{(0,0)\}\right)=1.$$

Niech τ będzie strategią wybraną przez studenta. Wówczas $\tau=n$ oznacza, że student będzie zdawał jako $(n+1)$-wszy. Zda egzamin gdy zajdzie zdarzenie

$$A=\{k_\tau-k_{\tau+1}=1\}.$$

Stąd

$$\begin{aligned}
\mathbb{P}(A) &= \sum_{n=0}^{K+M}\mathbb{P}\left(\tau=n \quad \text{i} \quad k_n-k_{n+1}=1\right)\\
&= \sum_{n=0}^{K+M}\mathbb{E}\left(\chi_{\{\tau=n\}}\mathbb{P}\left(k_n=k_{n+1}+1 \mid \sigma(X_0,\dots,X_n)\right)\right)\\
&= \sum_{n=0}^{K+M}\mathbb{E}\left(\chi_{\{\tau=n\}}\frac{k_n}{k_n+m_n}\right)\\
&= \mathbb{E}\, r(X_\tau),
\end{aligned}$$

gdzie

$$r(k,m)=\frac{k}{k+m},\quad (k,m)\in E,\qquad \frac{0}{0}=0.$$

Zauważmy, że

$$\begin{aligned}
Pr(k,m) &= P\left((k,m),\{(k-1,m)\}\right)\frac{k-1}{k-1+m}\\
&\qquad +P\left((k,m),\{(k,m-1)\}\right)\frac{k}{k+m-1}\\
&= \frac{k}{k+m}\cdot\frac{k-1}{k-1+m}+\frac{m}{k+m}\cdot\frac{k}{k+m-1}\\
&= \frac{k(k-1)+km}{(k+m)(k-1+m)}=\frac{k}{k+m}\\
&= r(k,m).
\end{aligned}$$

Czyli $Pr=r$. W postaci funkcjonału zysku $q=0$ i $\gamma=1$. Stąd mamy

$$Qr(k,m)=\max\left\{Pr(k,m),r(k,m)\right\}=r(k,m).$$

A więc maksymalne prawdopodobieństwo zdania wynosi

$$Q^{K+M}r(K_0,M_0)=r(K_0,M_0)=r(K,M).$$

Z drugiej strony gdyby studentowi zależało na znalezieniu strategii minimalizującej prawdopodobieństwo zdania to do policzenia miałby operator

$$\tilde{Q}h:=\min\{Ph,r\}.$$

Ale, ponieważ $Pr=r$, więc $\tilde{Q}r=r$. Czyli minimalne prawdopodobieństwo wynosiłoby również $r(K,M)$! Stąd prawdopodobieństwa minimalne i maksymalne zdania egzaminu są sobie równe. Czyli każda strategia daje to samo prawdopodobieństwa zdania $K/(K+M)$.

5.5 Pewne uogólnienie

Podamy uogólnienie Twierdzenia 5.1 o optymalnym sterowaniu na skończonym horyzoncie na przypadek rodziny Markowa $(X_n, \mathcal{F}_n)$. W tej sytuacji dopuszczamy Markowskie momenty zatrzymania względem filtracji $(\mathcal{F}_n)$, która może być większa od filtracji $(\mathcal{X}_n)$ generowaej przez proces (X_n). Potrzeba tego typu uogólnienia zostanie uzasadniona w rozdziale poświęconym problemowi rozregulowania.

Niech $(X_n, \mathcal{F}_n)$ będzie rodziną Markowa na E z operatorem przejścia P. Ustalmy skończony horyzont czasowy N. Niech $\varSigma_N$ będzie rodziną wszystkich momentów Markowa τ względem filtracji $(\mathcal{F}_n)$ spełniających $\mathbb{P}(\tau \leq N) = 1$. Celem jest znalezienie optymalnego $\tau \in \varSigma_N$ dla funkcjonał u kosztu (lub zysku)

$$J_N(\tau, X_0) = \mathbb{E}\left(\sum_{n=0}^{\tau-1} \gamma^n q(X_n) + \gamma^\tau r(X_\tau)\right).$$

Zakładamy, że $\gamma \geq 0$ oraz, że q i r są nieujemnymi mierzalnymi funkcjami na E.

Zdefiniujemy operator Bellmana Q działający z przestrzeni funkcji mierzalnych nieujemnych jak w przypadku szczególnym. Mianowicie jeżeli J_N jest kosztem, to

$$Q\psi(x) = \min\left\{q(x) + \gamma P\psi(x), r(x)\right\}, \qquad \psi \geq 0,\ x \in E.$$

Jeżeli J_N jest zyskiem to

$$Q\psi(x) = \max\left\{q(x) + \gamma P\psi(x), r(x)\right\}, \qquad \psi \geq 0,\ x \in E.$$

Mamy następujące uogólnienie Twierdzenia 5.1.

Twierdzenie 5.4 *Moment zatrzymania*

$$\widehat{\tau} = \inf\{n \leq N \colon Q^{N-n} r(X_n) = r(X_n)\}$$

jest optymalny oraz zachodzi

$$J_N(\widehat{\tau}, X_0) = \mathbb{E}\, Q^N r(X_0) \tag{5.1}$$

Dowód Przeprowadzimy dowód dla funkcjonału kosztu. Dla funkcjonału zysku dowód jest analogiczny. Pokażemy najpierw (5.1). Ustalmy moment Markowa $\tau \in \varSigma_N$. Wówczas

$$
\begin{aligned}
J_N(\tau, X_0) &= \mathbb{E}\left(\sum_{n=0}^{\tau-1} \gamma^n q(X_n) + \gamma^\tau r(X_\tau)\right) \\
&= \mathbb{E}\left(\sum_{n=0}^{\tau-1} \gamma^n q(X_n) + \gamma^\tau r(X_\tau)\right) \chi_{\{\tau=N\}} \\
&\quad + \mathbb{E}\left(\sum_{n=0}^{\tau-1} \gamma^n q(X_n) + \gamma^\tau r(X_\tau)\right) \chi_{\{\tau<N\}} \\
&= \mathbb{E}\left(\sum_{n=0}^{N-1} \gamma^n q(X_n) + \gamma^N r(X_N)\right) \chi_{\{\tau>N-1\}} \\
&\quad + \mathbb{E}\left(\sum_{n=0}^{\tau-1} \gamma^n q(X_n) + \gamma^\tau r(X_\tau)\right) \chi_{\{\tau\le N-1\}}.
\end{aligned}
$$

Ponieważ

$$
\{\tau > N-1\} \in \mathcal{F}_{N-1}
$$

mamy

$$
\begin{aligned}
&\mathbb{E}\left(\sum_{n=0}^{N-1} \gamma^n q(X_n) + \gamma^N r(X_N)\right) \chi_{\{\tau>N-1\}} \\
&= \mathbb{E}\,\mathbb{E}\left[\left(\sum_{n=0}^{N-1} \gamma^n q(X_n) + \gamma^N r(X_N)\right) \chi_{\{\tau>N-1\}} \middle| \mathcal{F}_{N-1}\right] \\
&= \mathbb{E}\,\mathbb{E}\left[\left(\sum_{n=0}^{N-1} \gamma^n q(X_n) + \gamma^N r(X_N)\right) \middle| \mathcal{F}_{N-1}\right] \chi_{\{\tau>N-1\}}.
\end{aligned}
$$

Stąd

$$
\begin{aligned}
&\mathbb{E}\left(\sum_{n=0}^{N-1} \gamma^n q(X_n) + \gamma^N r(X_N)\right) \chi_{\{\tau>N-1\}} \\
&= \mathbb{E}\left(\sum_{n=0}^{N-1} \gamma^n q(X_n) + \gamma^N Pr(X_{N-1})\right) \chi_{\{\tau>N-1\}} \\
&= \mathbb{E}\left(\sum_{n=0}^{N-2} \gamma^n q(X_n) + \gamma^{N-1}\left(q(X_{N-1}) + \gamma Pr(X_{N-1})\right)\right) \chi_{\{\tau>N-1\}} \\
&\ge \mathbb{E}\left(\sum_{n=0}^{N-2} \gamma^n q(X_n) + \gamma^{N-1} Qr(X_{N-1})\right) \chi_{\{\tau>N-1\}}.
\end{aligned}
$$

Oczywiście z definicji operatora Q mamy

$$
Qr(X_\tau) \le r(X_\tau)
$$

Stąd

$$J_N(\tau, X_0) \geq \mathbb{E}\left(\sum_{n=0}^{\tau\wedge(N-1)-1} \gamma^n q(X_n) + \gamma^{\tau\wedge(N-1)} Qr(X_{\tau\wedge(N-1)})\right).$$

Powtarzając tą procedure $j-1$ razy dla momentu τ zastąpionego odpowiednio przez

$$\tau \wedge (N-1), \quad \tau \wedge (N-2), \quad \ldots, \tau \wedge (N-j)$$

oraz r zastąpionego przez

$$Qr, \quad Q^2 r, \quad \ldots, Q^{j-1}$$

otrzymujemy oszacowanie

$$J_N(\tau, X_0) \geq \mathbb{E}\left(\sum_{n=0}^{\tau\wedge(N-j)-1} \gamma^n q(X_n) + \gamma^{\tau\wedge(N-j)} Q^j r(X_{\tau\wedge(N-j)})\right).$$

Biorąc teraz $j = N$ otrzymujemy (5.1) co dowodzi pierwszej części twierdzenia. Do dowodu optymalności $\widehat{\tau}$ wystarczy pokazać równość

$$J_N(\widehat{\tau}, X_0) = \mathbb{E}\, Q^N r(X_0).$$

Z definicji $\widehat{\tau}$ wynika, że jest to moment Markowa ze względu na filtrację $(\mathcal{X}_n)$, gdzie

$$\mathcal{X}_n = \sigma(X_0, X_1, \ldots, X_n), \qquad n = 0, 1, \ldots.$$

Tak więc wymagana równość wynika z Twierdzenia 5.1. □

6

Stopowanie - horyzont nieskończony

Nie zawsze istnieje strategia optymalna dla problemu stopowania na nieskończonym przedziale czasowym. Istotnie, niech $\{X_n\}$ będzie ciągiem niezależnych zmiennych losowych o rozkładzie jednostajnym na odcinku $[0,1]$. Wówczas

$$\sup_{\tau \in \Sigma} \mathbb{E}\, X_\tau \chi_{\{\tau<+\infty\}} = 1,$$

gdzie Σ oznacza zbiór wszystkich momentów Markowa względem filtracji zadanej przez $\{X_n\}$. Ale, ponieważ

$$\forall\, \tau \in \Sigma, \quad \mathbb{P}\left(X_\tau \chi_{\{\tau<+\infty\}} < 1\right) = 1,$$

więc

$$\forall\, \tau \in \Sigma, \quad \mathbb{E}\, X_\tau \chi_{\{\tau<+\infty\}} < 1.$$

6.1 Zasadnicze twierdzenia

Pierwszy rezultat związany jest z maksymalizacją funkcjonału zysku postaci

$$J(\tau, X_0) := \mathbb{E}\, \gamma^\tau r(X_\tau)\chi_{\{\tau<+\infty\}}, \qquad \tau \in \Sigma,$$

gdzie $r\colon E \to [0,+\infty)$, $\gamma > 0$, a $\{X_n\}$ zadany jest rekurencyjnie

$$X_{n+1} = F(X_n, \xi_n),$$

$\{\xi_n\}$ ciąg iid na $(S, \mathcal{S})$. Niech P będzie prawdopodobieństwem przejścia

$$Ph(x) := \mathbb{E}\, h(F(x,\xi_0)), \qquad x \in E.$$

Niech

$$Qh(x) := \max\{\gamma Ph(x), r(x)\},$$

oraz niech $v_n(x) = Q^n r(x)$. Wówczas, ciąg $\{v_n\}$ jest rosnący i jako taki zbieżny do funkcji v_∞ o wartościach w $[0,+\infty]$.

Twierdzenie 6.1 *Funkcja v_∞ jest minimalnym nieujemnym rozwiązaniem równania*

$$v(x) = Qv(x), \qquad x \in E.$$

Ponadto, jeżeli moment zatrzymania

$$\widehat{\tau} := \inf\{n \colon v_\infty(X_n) = r(X_n)\}$$

spełnia

$$\lim_{N\to+\infty} \gamma^N \mathbb{E}\, v_\infty(X_N)\chi_{\{\widehat{\tau}\geq N\}} = 0,$$

to maksymalizuje on problem

$$\sup_{\tau\in\Sigma} J(\tau, X_0),$$

oraz

$$\mathbb{E}\, v_\infty(X_0) = \sup_{\tau\in\Sigma} J(\tau, X_0) = J(\widehat{\tau}, X_0).$$

Dowód Ponieważ Q policzony na funkcji tożsamościowo równej 0 daje r, mamy

$$v_\infty(x) = \lim_{n\to+\infty} Q^n(0)(x), \qquad x \in E.$$

Stąd i z monotoniczności $v_\infty \leq Qv_\infty$. Ponieważ

$$Q^{n+1}(0)(x) \geq \max\{\gamma PQ^n(0)(x), r(x)\}$$

przechodząc do granicy otrzymujemy

$$v_\infty(x) \geq \max\{\gamma PQ^n(0)(x), r(x)\}$$

a następnie

$$v_\infty(x) \geq Qv_\infty(x).$$

Czyli $v_\infty = Qv_\infty$. Minimalność wynika z monotoniczności Q i z faktu że v_∞ jest granicą punktową ciągu $\{Q^n(0)\}$.

Niech

$$J_N(\tau, X_0) := \mathbb{E}\, \gamma^\tau r(X_\tau)\chi_{\{\tau\leq N\}}.$$

Wówczas, z twierdzenia o zbieżności monotonicznej,

$$J(\tau, X_0) = \lim_{N\to+\infty} J_N(\tau, X_0).$$

Dla J_N maksymalny zysk wynosi

$$\mathbb{E}\, Q^N r(X_0) \leq \mathbb{E}\, v_\infty(X_0).$$

Stąd

$$J(\tau, X_0) \leq \mathbb{E}\, v_\infty(X_0).$$

Niech $\widehat{\tau}$ będzie momentem zdefiniowanym w twierdzeniu. Wystarczy pokazać, że

$$J(\widehat{\tau}, X_0) = \mathbb{E}\, v_\infty(X_0).$$

W tym celu, zauważmy, że z definicji $\widehat{\tau}$,

$$r(X_{\widehat{\tau}})\chi_{\{\widehat{\tau}<+\infty\}} = v_\infty(X_{\widehat{\tau}})\chi_{\{\widehat{\tau}<+\infty\}}.$$

Stąd

$$\begin{aligned} J(\widehat{\tau}, X_0) &= \lim_{N\to+\infty} J_N(\widehat{\tau}, X_0) \\ &= \lim_{N\to\infty} \mathbb{E}\, \gamma^{\widehat{\tau}} r(X_{\widehat{\tau}})\chi_{\{\widehat{\tau}\le N\}} \\ &= \lim_{N\to+\infty} \mathbb{E}\, \gamma^{\widehat{\tau}} v_\infty(X_{\widehat{\tau}})\chi_{\{\widehat{\tau}\le N\}} \\ &= \lim_{N\to+\infty} \mathbb{E}\, \gamma^{\widehat{\tau}\wedge N} v_\infty(X_{\widehat{\tau}\wedge N})\chi_{\{\widehat{\tau}\le N\}}. \end{aligned}$$

Z założenia

$$\begin{aligned} &\lim_{N\to+\infty} \mathbb{E}\, \gamma^{\widehat{\tau}\wedge N} v_\infty(X_{\widehat{\tau}\wedge N})\chi_{\{\widehat{\tau}\le N\}} \\ &= \lim_{N\to+\infty} \mathbb{E}\, \left(\gamma^{\widehat{\tau}\wedge N} v_\infty(X_{\widehat{\tau}\wedge N})\chi_{\{\widehat{\tau}\le N\}} + \gamma^N v_\infty(X_N)\chi_{\{\widehat{\tau}>N\}}\right) \\ &= \lim_{N\to+\infty} \mathbb{E}\, \gamma^{\widehat{\tau}\wedge N} v_\infty(X_{\widehat{\tau}\wedge N}). \end{aligned}$$

Stąd

$$J(\widehat{\tau}, X_0) = \lim_{N\to+\infty} \mathbb{E}\, \gamma^{\widehat{\tau}\wedge N} v_\infty(X_{\widehat{\tau}\wedge N}).$$

Ustalmy skończony horyzont N. Dowód będzie zakończony gdy pokażemy, że

$$\mathbb{E}\, \gamma^{\widehat{\tau}\wedge N} v_\infty(X_{\widehat{\tau}\wedge N}) = \mathbb{E}\, v_\infty(X_0).$$

Oczywiście wystarczy pokazać, że dla dowolnego $N \ge 1$ zachodzi

$$\mathbb{E}\, \gamma^{\widehat{\tau}\wedge N} v_\infty(X_{\widehat{\tau}\wedge N}) = \mathbb{E}\, \gamma^{\widehat{\tau}\wedge(N-1)} v_\infty(X_{\widehat{\tau}\wedge(N-1)}).$$

Mamy

$$\begin{aligned} \mathbb{E}\, \gamma^{\widehat{\tau}\wedge N} v_\infty(X_{\widehat{\tau}\wedge N}) = \mathbb{E}\, &\gamma^{\widehat{\tau}\wedge(N-1)} v_\infty(X_{\widehat{\tau}\wedge(N-1)})\chi_{\{\widehat{\tau}\le(N-1)\}} \\ &+\mathbb{E}\, \gamma^{\widehat{\tau}\wedge N} v_\infty(X_{\widehat{\tau}\wedge N})\chi_{\{\widehat{\tau}>(N-1)\}}. \end{aligned}$$

Następnie

$$\begin{aligned} &\mathbb{E}\, \gamma^{\widehat{\tau}\wedge N} v_\infty(X_{\widehat{\tau}\wedge N})\chi_{\{\widehat{\tau}>(N-1)\}} \\ &= \mathbb{E}\,\mathbb{E}\left(\gamma^{\widehat{\tau}\wedge N} v_\infty(X_{\widehat{\tau}\wedge N})\chi_{\{\widehat{\tau}>(N-1)\}}|X_{N-1}\right) \\ &= \mathbb{E}\, \gamma^{N-1}\, \gamma\, \mathbb{E}\,(v_\infty(X_N)|X_{N-1})\ \chi_{\{\widehat{\tau}>(N-1)\}} \\ &= \mathbb{E}\, \gamma^{N-1}\, \gamma P v_\infty(X_{N-1})\ \chi_{\{\widehat{\tau}>(N-1)\}}. \end{aligned}$$

Tak więc wystarczy zauważyć, że ponieważ

$$v_\infty(x) = \max\left\{\gamma P v_\infty(x), r(x)\right\}, \qquad x \in E,$$

i

$$v_\infty(X_n) > r(X_n), \qquad n < \widehat{\tau},$$

zachodzi

$$\gamma P v_\infty(X_{N-1})\,\chi_{\{\widehat{\tau}>(N-1)\}} = v_\infty(X_{N-1})\,\chi_{\{\widehat{\tau}>(N-1)\}}.$$

□

Niech v_∞ będzie funkcją skonstruowaną w Twierdzeniu 6.1. Niech

$$\mathcal{C} = \left\{x \in E\colon v_\infty(x) = r(x)\right\}.$$

Wówczas

$$\widehat{\tau} = \inf\left\{n\colon X_n \in \mathcal{C}\right\}.$$

Zbiór $\mathcal{C}$ nazywamy *coincidence set* dla naszego problemu. W pewnych przypadkach może on być wyznaczony w sposób bezpośredni. Mamy następujący rezultat. Dotyczy on zagadnienia optymalnego stopowania dla funkcjonału zysku

$$J(\tau, X_0) = \mathbb{E}\, r(X_\tau)\chi_{\{\tau<+\infty\}},$$

gdzie (X_n) jest łańcucha Markowa o prawdopodobieństwie przejścia $P(x, \Gamma)$, $x \in E$, $\Gamma \in \Gamma$,

Twierdzenie 6.2 *Załóżmy, że r jest mierzalną, ograniczoną funkcją nieujemną na E. Niech*

$$\mathcal{C} = \left\{x \in E\colon Pr(x) \le r(x)\right\}.$$

Załóżmy, że dla każdego $x \in \mathcal{C}$, $P(x, \mathcal{C}) = 1$, oraz, że dla każdego $x \in E$,

$$\mathbb{P}\left(X_n \in \mathcal{C} \text{ dla pewnego } n = 0, 1, \ldots | X_0 = x\right) = 1.$$

Wówczas moment zatrzymania

$$\tau_{\mathcal{C}} = \inf\left\{n\colon X_n \in \mathcal{C}\right\}$$

jest optymalny.

Dowód Zdefiniujemy nowe prawdopodobieństwo przejścia $\tilde{P}$ na E kładąc

$$\tilde{P}(x, \Gamma) = P(x, \Gamma), \quad x \notin \mathcal{C}, \quad \Gamma \in \mathcal{E}$$

oraz

$$\tilde{P}(x, \{x\}) = 1, \quad x \in \mathcal{C}.$$

Wówczas $\tilde{P}$ jest prawdopodobienstwem przejścia łańcucha

$$\tilde{X}_n = X_{n\wedge\tau_{\mathcal{C}}}, \qquad n = 0, 1, \dots.$$

Ponadto

$$\tilde{P}^n r(x) \geq r(x), \qquad n = 0, 1, \dots, \quad x \in E.$$

Wynika to z faktu, że

$$\tilde{P}r(x) = Pr(x) > r(x) \qquad \text{dla} \quad x \notin \mathcal{C}.$$

Pokażemy, że funkcje

$$v_{n+1} := \max\left(r, Pv_n\right), \quad n = 0, 1, \dots, \quad v_0 = r$$

spełniają

$$v_n = \tilde{P}^n r.$$

Istotnie, powyższa równość zachodzi dla $n = 0$. Załóżmy, że zachodzi ona dla ustalonego n. Wówczas

$$Pv_n(x) = \tilde{P}v_n(x) = \tilde{P}^{n+1}r(x), \qquad \text{dla } x \notin \mathcal{C}.$$

Ponieważ $Pv_n \geq r$. Mamy więc

$$v_{n+1}(x) = Pv_n(x) = \tilde{P}^{n+1}r(x), \qquad x \notin \mathcal{C}.$$

Dla $x \in \mathcal{C}$ mamy

$$v_n(x) = \tilde{P}^n r(x) = r(x).$$

Stąd i z definicji zbioru $\mathcal{C}$,

$$\begin{aligned} v_{n+1}(x) &= \max\left\{Pv_n(x), r(x)\right\} = \max\left\{Pr(x), r(x)\right\} \\ &= r(x) = \tilde{P}^{n+1}r(x). \end{aligned}$$

Tak więc równość $v_n = \tilde{P}^n r$ jest wykazana. Ponieważ $\tilde{P}$ jest prawdopodobieństwem przejścia dla łańcucha $\{X_{n\wedge\tau_{\mathcal{C}}}\}$ mamy

$$\mathbb{E}\, v_N(X_0) = \mathbb{E}\, \tilde{P}^N r(X_0) = \mathbb{E}\, r(X_{N\wedge\tau_{\mathcal{C}}}).$$

Z twierdzenia dla skończonego horyzontu

$$\mathbb{E}\, v_N(X_0) \geq \mathbb{E}\, r(X_{\tau\wedge N}), \qquad \forall\, N, \quad \forall\, \tau \in \Sigma.$$

Stąd

$$\mathbb{E}\, r(X_{N\wedge\tau_{\mathcal{C}}}) \geq \mathbb{E}\, r(X_{\tau\wedge N}), \qquad \forall\, N, \quad \forall\, \tau \in \Sigma.$$

Przechodząc do granicy $N \to +\infty$ i korzystając z faktu że $\mathbb{P}\left(\tau_{\mathcal{C}} < +\infty\right) = 1$ otrzymujemy żądany wniosek

$$\mathbb{E}\, r(X_{\tau_C}) \geq \sup_{\tau \in \Sigma} \mathbb{E}\, r(X_\tau).$$

Oczywiście dla granicy

$$v_\infty(x) = \lim_{n \to +\infty} \tilde{P}^n r(x), \qquad x \in E$$

mamy

$$\mathbb{E}\, v_\infty(X_0) = \mathbb{E}\, r(X_{\tau_C}).$$

□

6.2 Zastosowania do błądzenia przypadkowego

Proces, który będziemy zatrzymywali jest symetrycznym błądzeniem przypadkowym na skończonym zbiorze $E = \{0, 1, \ldots, K\}$ z punktami pochłaniającymi 0 i K. Dokładnie prawdopodobieństwo przejścia dane jest przez macierz $P = (p_{i,j})$, $p_{i,j} = P(X_{n+1} = j | X_n = i)$. Mianowicie

$$p_{0,0} = p_{K,K} = 1 \qquad i \qquad p_{i,i-1} = \frac{1}{2} = p_{i,i+1}, \quad i = 1, \ldots, K-1.$$

Szukamy momentu Markowa $\widehat{\tau}$ takiego, że

$$\mathbb{E}\, r(X_{\widehat{\tau}}) \chi_{\{\widehat{\tau} < +\infty\}} \geq \sup_{\tau \in \Sigma} \mathbb{E}\, r(X_{\widehat{\tau}}) \chi_{\{\tau < +\infty\}},$$

gdzie r jest nieujemną funkcją na E, taką że $r(0) = 0 = r(K)$. Aby rozwiązać problem wystarczy zauważyć, że v_∞ jest najmniejszą funkcją nieujemną spełniającą:

1) $v_\infty(i) \geq r(i)$, $i \in E$,
2) $v_\infty(i) \geq P v_\infty(i) = \frac{1}{2} \left(v_\infty(i+1) + v_\infty(i-1) \right)$, $i = 1, \ldots, K-1$.

Tak więc v_∞ jest wklęsłą otoczką r. Ponieważ $v_\infty(0) = v_\infty(K) = 0$ oraz

$$\mathbb{P}\left(\lim_{n \to +\infty} X_n \in \{0, K\} \right) = 1,$$

mamy

$$\lim_{N \to +\infty} \mathbb{E}\, v_\infty(X_N) = 0.$$

Tak więc warunek w Twierdzeniu 6.1 jest spełniony i moment zatrzymania $\widehat{\tau}$ dany jest wzorem

$$\widehat{\tau} = \inf \left\{ n \colon v_\infty(X_n) = r(X_n) \right\}.$$

6.3 Rozwiązanie problemu sekretarki

Podamy rozwiązanie problemu sekretarki, patrz Przykład 1.3.

Niech $\{\xi_n\}$, $n = 1, \ldots, N$ będzie ciągiem niezależnych zmiennych losowych o rozkładzie jednostajnym na $[0, 1]$. Zdefiniujmy $\{X_n\}$ w następujący sposób: $X_1 = 1$,

$$X_{n+1} = \inf \{l > X_n \colon l \leq N, \quad \xi_l \geq \xi_{X_n}\}.$$

Tutaj przyjmujemy konwencje $\inf \emptyset = +\infty$.

Interpretujemy ξ_n jako kompetencje n-tej kandydatki. Kompetencja waha się od najniższej 0 do najwyższej 1. Zmienna losowa X_n, o ile $X_n \neq +\infty$, podaje numer pierwszej kandydatki lepszej od kandydatki do numeru X_{n-1} włącznie. Zdarzenie

$$X_n < +\infty, \quad X_{n+1} = +\infty$$

oznacza, że X_n-ta kandydatka miała najwyższe kwalifikacje spośród wszystkich N.

Osoba rekrutująca bada kwalifikacje kandydatki i na ich podstawie podejmuje decyzje. Ponieważ interesuje go jedynie wybór najlepszej, decyzje podejmuje w chwilach X_1, X_2,.... Pokażemy, że problem sekretarki jest równoważny problemowi optymalnego stopowania procesem (X_n) na przestrzeni stanów $E = \{1, \ldots, N, +\infty\}$, z funkcjonałem zysku

$$J(\tau, 1) = \mathbb{E}\, r(X_\tau)\chi_{\{\tau<+\infty\}}, \qquad \tau \in \Sigma,$$

gdzie

$$r(k) = p_{k,+\infty} = P\left(X_{n+1} = +\infty | X_n = k\right), \qquad k \leq N, \ r(+\infty) = 0.$$

Istotnie, zauważmy, że

$$\lim_{n\to+\infty} X_n = +\infty := X_\infty.$$

Ponieważ, $r(+\infty) = 0$ mamy

$$\mathbb{E}\, r(X_\tau)\chi_{\{\tau<\infty\}} = \mathbb{E}\, r(X_\tau).$$

Trzeba więc wykazać, że

$$\mathbb{P}\left(X_\tau \text{ jest najlepsza}\right) = \mathbb{E}\, r(X_\tau).$$

Wynika to z następującego rachunku

$$
\begin{aligned}
&\mathbb{P}\left(X_\tau \text{ jest najlepsza}\right)\\
&= \sum_{n=1}^{N} \mathbb{P}\left(X_n \text{ jest najlepsza i } \tau = n\right)\\
&= \sum_{n=1}^{N} \mathbb{E}\,\mathbb{P}\left(X_n \text{ jest najlepsza i } \tau = n | X_n\right)\\
&= \sum_{n=1}^{N} \mathbb{E}\,\chi_{\{\tau=n\}}\,\mathbb{P}\left(X_n \text{ jest najlepsza} \,| X_n\right)\\
&= \sum_{n=1}^{N} \mathbb{E}\,\chi_{\{\tau=n\}} \sum_{k=1}^{N} \mathbb{P}\left(X_n \text{ jest najlepsza} \,| X_n = k\right) \chi_{\{X_n=k\}}\\
&= \sum_{n=1}^{N} \mathbb{E}\,\chi_{\{\tau=n\}} \sum_{k=1}^{N} r(k)\chi_{\{X_n=k\}}\\
&= \sum_{n=1}^{N} \mathbb{E}\,\chi_{\{\tau=n\}} r(X_n)\\
&= \mathbb{E}\, r(X_\tau).
\end{aligned}
$$

Rozwiązanie problemu sekretarki zawarte jest w następującym twierdzeniu.

Twierdzenie 6.3 *Niech*

$$k_N = \inf\left\{k = 1, \ldots N \colon \frac{1}{k} + \frac{1}{k+1} + \ldots + \frac{1}{N-1} \le 1\right\}.$$

Wówczas moment

$$\widehat{\tau} = \inf\{n \colon X_n \ge k_N\}$$

jest optymalny.

W dowodzie twierdzenia wykorzystamy następujący rezultat.

Lemat 6.1 (X_n) *jest łańcuchem Markowa o prawdopodobieństwach przejściach*

$$p_{k,l} = \frac{k}{l(l-1)}, \quad k < l \le N,$$

$$p_{k,+\infty} = \frac{k}{N}, \quad k < +\infty, \qquad p_{+\infty,+\infty} = 1.$$

Dowód twierdzenia Stosujemy Twierdzenie 6.2. W tym celu znajdziemy "the coincidence set"

$$\mathcal{C} = \{k \colon Pr(k) \le r(k)\}\,.$$

Ponieważ $r(+\infty) = 0$ i $Pr(+\infty) = r(+\infty) = 0$ mamy $+\infty \in \mathcal{C}$. Niech $k \le N$. Wówczas

$$r(k) = \frac{k}{N}$$

oraz

$$Pr(k) = \sum_{l=k+1}^{N} r(l) p_{k,l} = \sum_{l=k+1}^{N} \frac{l}{N} \frac{k}{l(l-1)} = \sum_{l=k+1}^{N} \frac{k}{(l-1)N}.$$

Stąd

$$\begin{aligned} Pr(k) \leq r(k) \quad &\Leftrightarrow \quad \sum_{k+1}^{N} \frac{k}{(l-1)N} \leq \frac{k}{N} \\ &\Leftrightarrow \quad \sum_{k+1}^{N} \frac{1}{l-1} \leq 1. \end{aligned}$$

A więc

$$\begin{aligned} \mathcal{C} &= \left\{ k \colon \frac{1}{k} + \frac{1}{k+1} + \ldots + \frac{1}{N-1} \leq 1 \right\} \cup \{+\infty\} \\ &= \{k \colon k_N \leq k \leq N\} \cup \{+\infty\}. \end{aligned}$$

□

6.4 Problem wynajmu apartamentu

Przedstawiamy problem wynajmu apartamentu sformułowany w [8], patrz również [5] i [7].

Załóżmy, że mamy zadany czas $T \in (0, +\infty)$ na wynajęcie apartamentu. Zgłoszenia wpływają zgodnie z procesem Poissona z intensywnością λ. To znaczy η_1, η_2,... są niezależnymi zmiennymi o rozkładzie wykładniczym

$$\mathbb{P}(\eta_n \leq t) = 1 - e^{-\lambda t}, \qquad t \geq 0,\ n = 1, 2, \ldots,$$

a

$$\pi_1 = \eta_1, \qquad \pi_2 = \eta_1 + \eta_2, \ldots \quad \ldots \pi_n = \sum_{j=1}^{n} \eta_j$$

to kolejne momenty zgłoszeń. Po otrzymaniu zgłoszenia musimy natychmiast podjąć decyzję czy daną ofertę przyjąć czy też odrzucić i czekać na nową. Załóżmy, że jakość wszystkich oferowanych apartamentów jest jednakowa. Jednak ceny wynajmu są różne. Podobnie jak w problemie sekretarki, naszym celem jest znalezienie strategii, która maksymalizuje prawdopodobieństwo wyboru najtańszego apartamentu.

Niech $\{\xi_n\}$ będzie ciągiem niezależnych zmiennych losowych o jednakowym (dowolnym) rozkładzie z gęstością w $(0, +\infty)$. Wartość ξ_n interpretujemy jako cenę wynajmu apartamentu zgłoszonego w momencie π_n. Tak jak w przypadku problemu sekretarki decyzję podejmujemy tylko wtedy gdy zgłoszenie jest *zgłoszeniem istotnym*, to znaczy takim, że

$$\xi_n < \min\{\xi_j : j < n\}.$$

Niech

$$E = \{\delta\} \cup \{\partial\} \cup \{1, 2, 3, \ldots\} \times [0, T].$$

Zdefiniujmy proces (X_n) o wartościach w E w następujący sposób: $X_0 = \delta$, $X_n = (m, t)$ wtedy i tylko wtedy gdy

$$\xi_n < \min_{j<n} \{\xi_j\} \quad \text{i} \quad \eta_1 + \ldots \eta_n = T - t,$$

to znaczy n-te zgłoszenie było istotne. W tym przypadku m oznacza liczbę zgłoszeń istotnych poprzedzających n-te zgłoszenie. Jeżeli

$$\xi_n \geq \min_{j<n} \{\xi_j\} \quad \text{lub} \quad \eta_1 + \ldots \eta_n > T,$$

to przyjmujemy, że $X_n = \partial$. Mamy następujący fakt, którego dowód można znaleźć w [7].

Lemat 6.2 (X_n) *jest jednorodnym procesem Markowa z prawdopodobieństwem przejścia*

$$p(x, \{y\}) = \mathbb{P}\left(X_{n+1} = y | X_n = x\right)$$

danym wzorem

$$p(\delta, \{\partial\}) = e^{-\lambda T}, \quad p(\partial, \{\partial\}) = 1,$$

$$\begin{aligned} p(\delta, \{(m, t)\}) &= \lambda e^{-\lambda(T-t)} \;\; \textit{gdy} \;\; m = 1, \\ p(\delta\{(m, t)\}) &= 0 \qquad\qquad\quad \textit{gdy} \;\; m \geq 2, \end{aligned}$$

$$\begin{aligned} &p((m, t), \{(m + k, t - u)\}) \\ &= \begin{cases} \frac{m\lambda^k u^{k-1} e^{-\lambda u}}{(k-1)!(m+k-1)(m+k)} & \textit{gdy } k \geq 1,\ u \in (0, t], \\ 0 & \textit{w przeciwnym przypadku}, \end{cases} \end{aligned}$$

$$p((m, t)\{\partial\}) = m e^{-\lambda t} \sum_{j=0}^{+\infty} \frac{(\lambda t)^j}{j!(m + j)}.$$

Niech Σ oznacza zbiór wszystkich momentów Markowa τ względem (X_n), takich że $\mathbb{P}(\tau < +\infty) = 1$. Problem wynajmu apartamentu sprowadza się do znalezienia momentu Markowa $\widehat{\tau} \in \Sigma$, takiego że

$$\mathbb{E}\, r(X_{\widehat{\tau}}) \geq \mathbb{E}\, r(X_\tau), \qquad \tau \in \Sigma,$$

gdzie

$$r(m, t) = p((m, t), \{\partial\}).$$

Zachodzi następujące twierdzenie.

Twierdzenie 6.4 *Optymalny moment $\widehat{\tau}$ dla problemu wynajmu apartamentu dany jest wzorem*

$$\widehat{\tau} = \inf\{n\colon X_n = (m,t) \text{ i } \lambda t \le x_m\}$$

gdzie x_m jest jedynym rozwiązaniem równania

$$\sum_{j=0}^{+\infty} \frac{x^j}{j!(m+j)} = \sum_{j=1}^{+\infty} \frac{x^j}{j!(m+j)} \sum_{k=1}^{j} \frac{1}{k+m-1}, \qquad x \ge 0.$$

Szkic dowódu Skorzystamy z Twierdzenia 6.2 wyznaczając "coincidence set". Niech P będzie operatorem przejścia. Z lematu mamy

$$Pr(m,t) = me^{-\lambda t} \sum_{j=1}^{+\infty} x^n a_{jm} b_{jm},$$

gdzie

$$x = \lambda t, \qquad a_{jm} = \frac{1}{j!(m+j)}, \qquad b_{jm} = \sum_{k=1}^{j} \frac{1}{k+m-1}.$$

Stąd

$$Pr(m,t) \le r(m,t)$$

wtedy i tylko wtedy gdy

$$\sum_{j=1}^{+\infty} x^m a_{jm} b_{jm} \le \sum_{j=0}^{+\infty} x^j a_{jm}.$$

Ostatecznie teza twierdzenia wynika z następujacego analitycznego lematu, którego dowód można znaleść w [8].

Lemat 6.3 *Niech*

$$h_m(x) = a_{0m} + \sum_{j=1}^{+\infty} (1 - b_{jm}) a_{jm} x^j, \qquad m \ge 1,\ x > 0.$$

Wówczas dla dowolnego $m \ge 1$ istnieje dokładnie jedno $x_m > 0$, takie że $h_m(x_m) = 0$. Ciąg (x_m) jest rosnący, ponadto, jeżeli $0 < x < x_m$, to $h_m(x) > 0$. Dla $x > x_m$ zachodzi $h_m(x) < 0$.

6.5 Stopowanie z ceną za zwłokę

Niech (X_n) będzie łańcuchem Markowa z operatorem przejścia P. Naszym celem będzie znalezienie momentu zatrzymania dla funkcjonału zysku

$$J(\tau, X_0) = \mathbb{E}\left(\gamma^\tau r(X_\tau) - \sum_{n=0}^{\tau-1} \gamma^n q(X_n)\right).$$

Będziemy zakładali, że $\gamma \in (0,1)$ oraz, że r i q są ograniczonymi, mierzalnymi i nieujemnymi funkcjami na przestrzeni stanów E. Ponieważ r jest ograniczona i $\gamma < 1$, przyjmujemy $\gamma^\tau r(X_\tau) = 0$ na zbiorze $\{\tau = +\infty\}$.

Wyrażenie $\gamma^n q(X_n)$ interpretujemy jako koszt za zwłoke ponoszony w momencie n, a $\gamma^\tau r(X_\tau)$ interpretujemy jako nagrode.

Twierdzenie 6.5 *(i) Istnieje dokładnie jedno ograniczone rozwiązanie V_∞ równania*

$$v(x) = \max\left(\gamma P v(x) - q(x), r(x)\right), \qquad x \in E. \tag{6.1}$$

(ii) Moment Markowa

$$\widehat{\tau} = \inf\{n\colon V_\infty(X_n) = r(X_n)\} \tag{6.2}$$

jest optymalny, to znaczy maksymalizuje funkcjonał zysku J.
(iii) Maksymalny zysk wynosi $\mathbb{E}\, V_\infty(X_0)$.

Dowód Pokażemy, że dla dowolnego momentu Markowa τ,

$$\mathbb{E}\sum_{n=0}^{\tau-1} \gamma^n q(X_n) = -\mathbb{E}\,\gamma^\tau H(X_\tau) + \mathbb{E}\, H(X_0), \tag{6.3}$$

gdzie

$$H(x) = \sum_{n=0}^{+\infty} \gamma^n P^n q(x), \qquad x \in E.$$

Dowód (6.3) podamy na końcu dowodu twierdzenia. Zakładając (6.3) otrzymujemy

$$\begin{aligned} J(\tau, X_0) &= \mathbb{E}\left(\gamma^\tau r(X_\tau) - \sum_{n=0}^{\tau-1} \gamma^n q(X_n)\right) \\ &= \mathbb{E}\left(\gamma^\tau (r+H)(X_\tau) - H(X_0)\right) \\ &= \mathbb{E}\,\gamma^\tau (r+H)(X_\tau) - \mathbb{E}\, H(X_0). \end{aligned}$$

Tak więc mamy teraz typowy problem stopowania z nieujemną funkcją $\tilde{r} = r + H$. Aby go rozwiązać szukamy minimalnego nieujemnego rozwiązania równania

$$w(x) = \max\left(\gamma P w(x), r(x) + H(x)\right), \qquad x \in E. \tag{6.4}$$

Załóżmy, że W_∞ jest tym rozwiązaniem. Wówczas W_∞ jest ograniczone. Istotnie niech $||q||_\infty$ i $||r||_\infty$ będą normami supremum q i r. Wówczas dla $x \in E$,

$$|H(x)| \le \sum_{n=0}^{+\infty} \gamma^n P^n q(x) \le \sum_{n=0}^{+\infty} \gamma^n ||q||_\infty \le \frac{||q||_\infty}{1-\gamma}.$$

Stąd H jest ograniczona. Funkcja W_∞ jest granicą punktową ciągu

$$W_{n+1}(x) = \max\left(\gamma P W_n(x), r(x) + H(x)\right), \quad W_0(x) = 0, \qquad x \in E.$$

Stąd, ponieważ $\gamma < 1$ i P jest kontrakcją mamy

$$||W_n||_\infty \leq ||r||_\infty + ||H||_\infty, \qquad n = 0, 1, \ldots.$$

Niech

$$\widehat{\tau} = \inf\left\{n \colon W_\infty(X_n) = r(X_n) + H(X_n)\right\}. \tag{6.5}$$

Z ograniczoności W_∞ i z faktu, że $\gamma < 1$ wynika

$$\lim_{N\to+\infty} \mathbb{E}\,\gamma^N W_\infty(X_N)\chi_{\{\widehat{\tau}\geq N\}} = 0.$$

Stąd $\widehat{\tau}$ jest optymalnym momentem zatrzymania dla wyjściowego funkcjonału J. Pokażemy, że

$$V_\infty = W_\infty - H$$

spełnia (6.1). Istotnie

$$\gamma P V_\infty(x) - q(x) = \gamma P W_\infty(x) - \gamma P H(x) - q(x).$$

Ale

$$\gamma P H(x) + q(x) = q(x) + \sum_{n=0}^{+\infty} \gamma^n P^{n+1} q(x) = H(x).$$

Stąd

$$\gamma P V_\infty(x) - q(x) = \gamma P W_\infty(x) - H(x), \qquad x \in E.$$

Czyli wystarczy zauważyć, że ponieważ W_∞ jest rozwiązaniem (6.4), mamy

$$W_\infty(x) - H(x) = \max\left(\gamma P W_\infty(x) - H(x), r(x)\right).$$

Czyli V_∞ jest rozwiązaniem (6.3). Oczywiście jest nieujemne bo

$$W_\infty \geq r + H$$

oraz ograniczone bo H i W_∞ są ograniczone. Ponadto $\widehat{\tau}$ zdefiniowany formułą (6.5) spełnia (6.2). Pokażemy, że V_∞ jest jedynym nieujemnym i ograniczonym rozwiązaniem (6.1). To oczywiście implikuje minimalność V_∞. Aby pokazać jedyność ograniczonego rozwiązanie (6.1) wystarczy pokazać, że operator

$$Q\psi(x) = \max\left(\gamma P\psi(x) - q(x), r(x)\right), \qquad x \in E$$

określony na przestrzeni $B_b(E)$ funkcji mierzalnych i ograniczonych, jest operatorem mocnej kontrakcji, to jest

$$||Q\psi - Q\varphi||_\infty \leq \gamma||\psi - \varphi||_\infty, \qquad \psi, \varphi \in B_b(E).$$

Wynika to z faktu, że dla dowolnych $\psi, \varphi \in B_b(E)$ oraz $x \in E$ mamy

$$
\begin{aligned}
&|Q\psi(x) - Q\varphi(x)| \\
&= |\max(\gamma P\psi(x) - q(x), r(x)) - \max(\gamma P\varphi(x) - q(x), r(x))| \\
&\leq \gamma\, |P\psi(x) - P\varphi(x)|\,.
\end{aligned}
$$

Tutaj wykorzystaliśmy elementarny fakt

$$
\left|\sup_\alpha a_\alpha - \sup_\alpha b_\alpha\right| \leq \sup_\alpha |a_\alpha - b_\alpha|\,.
$$

Pozostaje do wykazania (6.3). Wynika to z następującego rozumowania

$$
\mathbb{E} \sum_{n\geq\tau} \gamma^n q(X_n) = \mathbb{E} \sum_{k=0}^{+\infty} \sum_{n\geq k} \gamma^n q(X_n)\chi_{\{\tau=k\}}.
$$

Ponieważ dla $n \geq k$,

$$
\begin{aligned}
\mathbb{E}\,\gamma^n q(X_n)\chi_{\{\tau=k\}} &= \mathbb{E}\left(\mathbb{E}\left(\gamma^n q(X_n)\chi_{\{\tau=k\}}|X_k\right)\right) \\
&= \mathbb{E}\, P^{n-k} q(X_k)\chi_{\{\tau=k\}},
\end{aligned}
$$

mamy

$$
\begin{aligned}
\mathbb{E} \sum_{n\geq\tau} \gamma^n q(X_n) &= \mathbb{E} \sum_{k=0}^{+\infty} \sum_{n\geq k} \gamma^n P^{n-k} q(X_k)\chi_{\{\tau=k\}} \\
&= \mathbb{E} \sum_{k=0}^{+\infty} \gamma^k \left(\sum_{n\geq k} \gamma^{n-k} P^{n-k} q(X_k)\right) \chi_{\{\tau=k\}} \\
&= \mathbb{E} \sum_{k=0}^{+\infty} \gamma^k H(X_k)\chi_{\{\tau=k\}} \\
&= \mathbb{E}\,\gamma^\tau H(X_\tau).
\end{aligned}
$$

□

7

Obwiednia Snella

7.0.1 Istotny kres górny rodziny zmiennych losowych

Niech $\mathcal{I}$ będzie rodziną zmiennych losowych określonych na przestrzeni probabilistycznej $(\Omega, \mathfrak{F}, \mathbb{P})$. Przypomnijmy, że zmienne losowe to mierzalne odwzorowania z Ω w $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$. Jeśli $\mathcal{I}$ jest przeliczalna to

$$\sup_{\xi \in \mathcal{I}} \xi = \sup \{\xi \colon \xi \in \mathcal{I}\}$$

jest zmienną losową. Jeśli $\mathcal{I}$ jest nieprzeliczalna, to supremum nie musi być odwzorowaniem mierzalnym.

Przyklad 7.1 *Niech $\Omega = [0,1]$, $\mathfrak{F} = \mathcal{B}([0,1])$, a $\mathbb{P}$ niech będzie miarą Lebesgue'a na $[0,1]$. Rozważmy rodzinę*

$$\mathcal{I} = \left\{\chi_{\{x\}} \colon x \in \Gamma\right\},$$

gdzie $\Gamma \subset \mathbb{R}$ nie jest zbiorem mierzalnym. Oczywiście

$$\sup_{\xi \in \mathcal{I}} \xi = \sup_{x \in \Gamma} \chi_{\{x\}} = \chi_{\Gamma}.$$

Twierdzenie 7.1 *Niech $\mathcal{I}$ będzie rodziną zmiennych losowych określonych na przestrzeni probabilistycznej $(\Omega, \mathfrak{F}, \mathbb{P})$. Istnieje wówczas zmienna losowa η taka, że:*

(i) dla każdej $\xi \in \mathcal{I}$ zachodzi $\xi \leq \eta$, $\mathbb{P}$-p.n.
(ii)η jest najmniejszą (w sensie nierówności $\mathbb{P}$-p.n.) zmienną losową dla której zachodzi (i), to znaczy jeżeli $\tilde{\eta}$ jest zmienna losową taką, że $\xi \leq \tilde{\eta}$, $\mathbb{P}$-p.n. dla każdej $\xi \in \mathcal{I}$, to $\tilde{\eta} \leq \eta$.

Ponadto istnieje ciąg $(\xi_n) \subset \mathcal{I}$, taki, że

$$\eta = \sup_{n \in \mathbb{N}} \xi_n, \tag{7.1}$$

a jeżeli $\mathcal{I}$ jest rodziną filtrującą w górę, to znaczy, jeżeli dla dowolnych $\xi_1, \xi_2 \in \mathcal{I}$ istnieje $\xi \in \mathcal{I}$ taka, że $\xi_1 \vee \xi_2 \leq \xi$, $\mathbb{P}$-p.n. to istnieje ciąg $(\xi_n) \subset \mathcal{I}$, taki, że $\xi_n \uparrow \eta$, $\mathbb{P}$-p.n. Stąd, gdy rodzina $\mathcal{I}$ jest filtrująca w górę oraz

$$\mathbb{E}\,\xi^- < +\infty, \qquad \forall \xi \in I,$$

to zachodzi

$$\mathbb{E}\,\eta = \lim_{n\to\infty} \mathbb{E}\,\xi_n = \sup_{n\in\mathbb{N}} \mathbb{E}\,\xi_n = \sup_{\xi\in\mathcal{I}} \mathbb{E}\,\xi. \tag{7.2}$$

Zmienną losową η nazywamy *istotnym kresem górnym rodziny* $\mathcal{I}$ i oznaczamy $\operatorname{ess\,sup}_{\xi\in\mathcal{I}} \xi$. Oczywiście $\operatorname{ess\,sup}_{\xi\in\mathcal{I}} \xi$ jest wyznaczony z dokładnością do zbioru miary 0.

Dowód twierdzenia Bez straty ogólności możemy założyć, że wszystkie zmienne losowe z $\mathcal{I}$ przyjmują wartości z przedziału $[0,1]$. Istotnie istnieje ciągła rosnąca bijekcja z odcinka $[0,1]$ w $\overline{\mathbb{R}}$.

Niech $\mathcal{D}$ oznacza rodzinę wszystkich przeliczalnych podzbiorów $\mathcal{I}$. Dla $D \in \mathcal{D}$ niech

$$\xi_D = \sup_{\xi\in D} \xi.$$

Oczywiście, dla dowolnego $D \in \mathcal{D}$, ξ_D jest dobrze zdefiniowaną zmienną losową. Niech

$$a := \sup_{D\in\mathcal{D}} \mathbb{E}\,\xi_D.$$

Oczywiście $a \in [0,1]$ oraz istnieje ciąg $(D_n) \subset \mathcal{D}$ taki, że

$$\xi_{D_n} \uparrow a.$$

Niech $\widehat{D} = \bigcup_{n\in\mathbb{N}} D_n$. Wówczas

$$a = \mathbb{E}\,\xi_{\widehat{D}}.$$

Istotnie

$$\mathbb{E}\,\xi_{D_n} \leq \mathbb{E}\,\xi_{\widehat{D}} \leq a.$$

Pokażemy, że $\eta = \xi_{\widehat{D}}$ ma żądane własności. Z definicji zachodzi (7.1). Pokażemy, że dla dowolnej $\xi \in \mathcal{I}$, $\xi \leq \xi_{\widehat{D}}$, $\mathbb{P}$-p.n. Istotnie gdyby dla jakiegoś $\xi \in \mathcal{I}$, na zbiorze miary niezerowej zachodziłoby $\xi > \xi_{\widehat{D}}$, to $\xi \vee \xi_{\widehat{D}} \geq \xi_{\widehat{D}}$ oraz $\xi \vee \xi_{\widehat{D}} > \xi_{\widehat{D}}$ na zbiorze miary niezerowej. Stąd

$$\mathbb{E}\,\xi \vee \xi_{\widehat{D}} > \mathbb{E}\,\xi_{\widehat{D}},$$

co prowadzi do sprzeczności bo $\widehat{D} \cup \{\xi\} \in \mathcal{D}$. Oczywiście $\xi_{\widehat{D}}$ jest minimalny, bo gdy $\tilde{\eta}$ jest takie, że dla dowolnej $\xi \in \mathcal{I}$, $\xi \leq \tilde{\eta}$, $\mathbb{P}$-p.n., to $\xi_{\widehat{D}} \leq \tilde{\eta}$, $\mathbb{P}$-p.n.

Ustawmy teraz zbiór przeliczalny $\widehat{D}$ w ciąg (ξ_n). Wówczas

$$\xi_{\widehat{D}} = \sup_{n\in\mathbb{N}} \xi_n.$$

Ponadto, jeżeli rodzina $\mathcal{I}$ jest filtrowana w górę, to istnieje ciąg $(\tilde{\xi}_n) \subset \mathcal{I}$, taki, że $\tilde{\xi}_{n+1} \geq \xi_n \vee \tilde{\xi}_n$. Stąd

$$\tilde{\xi}_n \uparrow \xi_{\widehat{D}}, \qquad \mathbb{P} - \text{p.n.} \qquad \square$$

Zadanie 7.1 Pokazać, że dla rodziny z Przykładu 7.1, $\operatorname{ess\,sup}_{\xi \in \mathcal{I}} \xi = 0$.

7.0.2 Zasadnicze twierdzenie

Niech (Z_n) będzie ciągiem zmiennych losowych określonych na przestrzeni probabilistycznej $(\Omega, \mathfrak{F}, \mathbb{P})$. Inaczej (Z_n) jest procesem stochastycznym w czasie dyskretnym określonym na $(\Omega, \mathfrak{F}, \mathbb{P})$. Zakładamy, że $\mathbb{N} = \{0, 1, \ldots\}$. Niech $(\mathfrak{F}_n)$ będzie filtracją. Niech Σ oznacza ogół momentów Markowa τ względem filtracji $(\mathfrak{F}_n)$ takich, że $\tau < +\infty$, $\mathbb{P}$-p.n. Elementy Σ będziemy nazywali *regułami stopu.*

Załóżmy, że:

(i) dla dowolnego n, zmienna losowa Z_n jest $\mathfrak{F}_n$-mierzalna,
(ii)dla dowolnego n, zmienna losowa $Z_n \in L^1(\Omega, \mathfrak{F}_n, \mathbb{P})$,
$(iii)$$\sup_{n \in \mathbb{N}} Z_n^+ \in L^1(\Omega, \mathfrak{F}, \mathbb{P})$,
$(iv)$$Z_\tau^- \in L^1(\Omega, \mathfrak{F}, \mathbb{P})$ dla każdego $\tau \in \Sigma$.

Założenia (i), (ii) są naturalne. Założenia (iii) oraz (iv) są techniczne.

Definicja 7.1 Reguła stopu $\widehat{\tau} \in \Sigma$ jest optymalna gdy

$$\mathbb{E}\, Z_{\widehat{\tau}} = \sup_{\tau \in \Sigma} \mathbb{E}\, Z_\tau.$$

W celu znalezienia optymalnej reguły zatrzymania zdefinujemy tak zwaną *obwiednie Snella* (Y_n), procesu (Z_n) kładąc

$$Y_n := \operatorname*{ess\,sup}_{\tau \in \Sigma^n} \mathbb{E}\left(Z_\tau | \mathfrak{F}_n\right), \tag{7.3}$$

gdzie

$$\Sigma^n := \left\{\tau \in \Sigma \colon \tau \geq n,\ \mathbb{P} - \text{p.n.}\right\}.$$

Uwaga 7.1 Moment Markowa τ jest z klasy Σ^n wtedy i tylko wtedy gdy $\tau \geq n$, $\mathbb{P}$-p.n. Niech $\tilde{\Sigma}^n$ oznacza klasę wszystkich momentów Markowa takich, że $\tau \geq n$. Zauważmy, że dla $\tau \in \Sigma^n$, $\tau \vee n \in \tilde{\Sigma}^n$. Ponadto $\tau = \tau \vee n$, $\mathbb{P}$-p.n. Stąd

$$Y_n = \operatorname*{ess\,sup}_{\tau \in \Sigma^n} \mathbb{E}\left(Z_\tau | \mathfrak{F}_n\right) = \operatorname*{ess\,sup}_{\tau \in \tilde{\Sigma}^n} \mathbb{E}\left(Z_\tau | \mathfrak{F}_n\right). \tag{7.4}$$

Twierdzenie 7.2 *Proces* $Y := (Y_n)$ *zdefiniowany przez* (7.3) *ma następujące własności:*

(a) *Dla każdego* n, $Y_n \in L^1(\Omega, \mathfrak{F}, \mathbb{P})$, *czyli proces* Y *jest całkowalny.*
(b) *Dla każdego* n, $Y_n = \max\left\{Z_n, \mathbb{E}\left(Y_{n+1} | \mathfrak{F}_\mathfrak{n}\right)\right\}$.

(c) *Dla każdego* n, $\mathbb{E}\, Y_n = \sup_{\tau \in \Sigma^n} \mathbb{E}\, Z_\tau$.
(d) *Jeżeli* $\inf_{n\in\mathbb{N}} Z_n \in L^1(\Omega, \mathfrak{F}, \mathbb{P})$, *to* $(Y_n, \mathfrak{F}_\mathfrak{n})$ *jest najmniejszym nadmartyngałem majoryzującym proces* (Z_n).
(e) *Optymalna reguła zatrzymania istnieje wtedy i tylko wtedy gdy moment Markowa*

$$\tau_0 := \inf\{n \in \mathbb{N} \colon Y_n = Z_n\}$$

jest skończony $\mathbb{P}$*-p.n. Co więcej, jeżeli* $\mathbb{P}\{\tau_0 < +\infty\} = 1$, *to* τ_0 *jest najmniejszą optymalną regułą zatrzymania, to znaczy dla dowolnej optymalnej reguły zatrzymania* $\widehat{\tau}$ *zachodzi* $\widehat{\tau} \geq \tau_0$, $\mathbb{P}$*-p.n.*

7.0.3 Dowód zasadniczego twierdzenia

Lemat 7.1 *Dla dowolnego* n, *rodziana zmiennych losowych* $\mathbb{E}\,(Z_\tau|\mathfrak{F}_n)$, $\tau \in \Sigma^n$, *jest filtrująca w górę.*

Dowód Niech $\tau_1, \tau_2 \in \Sigma^n$. Niech

$$\begin{aligned} A &:= \{\omega \in \Omega \colon \mathbb{E}\,(Z_{\tau_1}|\mathfrak{F}_n)(\omega) \geq \mathbb{E}\,(Z_{\tau_2}|\mathfrak{F}_n)(\omega)\}, \\ B &:= \{\omega \in \Omega \colon \tau_1(\omega) \geq n \quad \text{i} \quad \tau_2(\omega) \geq n\}. \end{aligned}$$

Zdefiniujmy τ w następujący sposób:

$$\tau := \begin{cases} \tau_1 & \text{na zbiorze } A \cap B, \\ \tau_2 & \text{na zbiorze } A^c \cap B, \\ n & \text{na zbiorze } B^c. \end{cases}$$

Tutaj $\tau_i \geq n$, $i = 1, 2$ tylko na zbiorze miary 1. Tak więc τ musi być dodefiniowane na zbiorze B^c miary $\mathbb{P}$-zero. Zauważmy, że $\tau(\omega) \geq n$ dla wszystkich n. Następnie $A, B \in \mathfrak{F}_n$. Stąd

$$\{\tau = k\} = \begin{cases} \emptyset & \text{dla } k < n, \\ \{\tau_1 = k\} \cap A \cap B \cup \{\tau_2 = k\} \cap A^c \cap B & \text{dla } k > n, \\ B^c \cup \{\tau_1 = n\} \cap A \cap B \cup \{\tau_2 = n\} \cap A^c \cap B & \text{dla } k = n. \end{cases}$$

Stąd $\{\tau = k\} \in \mathfrak{F}_k$ dla każdego k, a więc τ jest momentem Markowa względem filtracji $(\mathfrak{F}_k)$. Ponieważ $\tau \geq n$, więc $\tau \in \Sigma^n$. Ponieważ $A, B \in \mathfrak{F}_n$, więc

$$\mathbb{E}\,(Z_\tau|\mathfrak{F}_n) \geq \mathbb{E}\,(Z_{\tau_1}|\mathfrak{F}_n) \vee \mathbb{E}\,(Z_{\tau_2}|\mathfrak{F}_n). \qquad \Box$$

Lemat 7.2 *Niech* $\tau \in \Sigma$. *Niech* Σ^τ *oznacza zbiór wszystkich momentów Markowa* σ *względem* $(\mathfrak{F}_k)$ *takich, że* $\sigma \geq \tau$, $\mathbb{P}$*-p.n. Wówczas rodzina zmiennych losowych*

$$\mathbb{E}\,(Z_\sigma|\mathfrak{F}_\tau), \qquad \sigma \in \Sigma^\tau,$$

jest filtrująca w górę.

Dowód Niech $\sigma_1, \sigma_2 \in \Sigma^\tau$. Niech

$$A := \{\omega \in \Omega \colon \mathbb{E}\left(Z_{\sigma_1} | \mathfrak{F}_\tau\right)(\omega) \geq \mathbb{E}\left(Z_{\sigma_2} | \mathfrak{F}_\tau\right)(\omega)\},$$
$$B := \{\omega \in \Omega \colon \sigma_1(\omega) \geq \tau(\omega) \quad \text{i} \quad \sigma_2(\omega) \geq \tau(\omega)\}.$$

Zdefiniujmy σ w następujący sposób:

$$\sigma := \begin{cases} \sigma_1 & \text{na zbiorze } A \cap B, \\ \sigma_2 & \text{na zbiorze } A^c \cap B, \\ \tau & \text{na zbiorze } B^c. \end{cases}$$

Mamy $A, B \in \mathfrak{F}_\tau$. Ponadto $\sigma(\omega) \geq \tau(\omega)$ dla wszystkich $\omega \in \Omega$. Pokazujemy, że σ jest momentem Markowa względem filtracji $(\mathfrak{F}_k)$. Mamy

$$\begin{aligned}\{\sigma = k\} &= \{\sigma_1 = k\} \cap A \cap B \cup \{\sigma_2 = k\} \cap A^c \cap B \cup \{\tau = k\} \cap B^c \\ &= \bigcup_{j=0}^{k} \{\sigma_1 = k\} \cap \{\tau = j\} \cap A \cap \{\sigma_2 \geq j\} \\ &\quad \cup \bigcup_{j=0}^{k} \{\sigma_2 = k\} \cap \{\tau = j\} \cap A^c \cap \{\sigma_1 \geq j\} \cup \{\tau = k\} \cap B^c.\end{aligned}$$

Ponieważ $B \in \mathfrak{F}_\tau$ więc $\{\tau = k\} \cap B^c \in \mathfrak{F}_k$. Podobnie, ponieważ $A \in \mathfrak{F}_\tau$, więc $A \cap \{\tau = j\} \in \mathfrak{F}_j$ i $A^c \cap \{\tau = j\} \in \mathfrak{F}_j \in \mathfrak{F}_j$. Stąd $\{\sigma = k\} \in \mathfrak{F}_k$. Czyli $\sigma \in \Sigma^\tau$.

Ponieważ $A, B \in \mathfrak{F}_\tau$ i $\mathbb{P}(B^c) = 0$, więc

$$\begin{aligned}\mathbb{E}\left(Z_\sigma | \mathfrak{F}_\tau\right) &= \mathbb{E}\left(Z_\sigma | \mathfrak{F}_\tau\right) \chi_{A \cap B} + \mathbb{E}\left(Z_\sigma | \mathfrak{F}_\tau\right) \chi_{A^c \cap B} \qquad \mathbb{P} - \text{p.n.} \\ &= \mathbb{E}\left(Z_{\sigma_1} | \mathfrak{F}_\tau\right) \chi_{A \cap B} + \mathbb{E}\left(Z_{\sigma_2} | \mathfrak{F}_\tau\right) \chi_{A^c \cap B} \qquad \mathbb{P} - \text{p.n.} \\ &\geq \mathbb{E}\left(Z_{\sigma_1} | \mathfrak{F}_\tau\right) \vee \mathbb{E}\left(Z_{\sigma_2} | \mathfrak{F}_\tau\right), \qquad \mathbb{P} - \text{p.n.} \qquad \square\end{aligned}$$

Dowód Twierdzenia (części (a), (b), (c), (d)) Przypomnijmy, patrz (7.3), że

$$Y_n := \operatorname*{ess\,sup}_{\tau \in \Sigma^n} \mathbb{E}\left(Z_\tau | \mathfrak{F}_n\right), \qquad n \in \mathbb{N}.$$

Pokażemy teraz punkt (a) twierdzenia, czyli, że $Y_n \in L^1(\Omega, \mathfrak{F}, \mathbb{P})$ dla dowolnego n. Ponieważ $\tau \equiv n \in \Sigma^n$ więc $Z_n \leq Y_n$, $\mathbb{P}$-p.n. Z drugiej strony

$$\operatorname*{ess\,sup}_{\tau \in \Sigma^n} Z_\tau \leq \sup_{k \geq n} Z_k^+ \in L^1(\Omega, \mathfrak{F}, \mathbb{P}).$$

Stąd $Y_n \leq \mathbb{E}\left(\sup_{k \geq n} Z_k^+ | \mathfrak{F}_n\right)$, $\mathbb{P}$-p.n. Tak więc z całkowalności Z_n i $\sup_{k \geq n} Z_k^+$ wynika całkowalność Y_n.

Pokażemy teraz część (b). Z Twierdzenia 7.1 i Lematu 7.1 wynika, że dla dowolnego n istnieje ciąg $(\tau_k) \subset \Sigma^{n+1}$ taki, że

$$\mathbb{E}\left(Z_{\tau_k} | \mathfrak{F}_{n+1}\right) \uparrow Y_{n+1}, \qquad \mathbb{P} - \text{p.n.}$$

Stąd
$$\mathbb{E}\left(\mathbb{E}\left(Z_{\tau_k}|\mathfrak{F}_{n+1}\right)|\mathfrak{F}_n\right) \uparrow \mathbb{E}\left(Y_{n+1}|\mathfrak{F}_n\right), \qquad \mathbb{P}-\text{p.n.}$$
Ponieważ, $\Sigma^{n+1} \subset \Sigma^n$, więc
$$\mathbb{E}\left(\mathbb{E}\left(Z_{\tau_k}|\mathfrak{F}_{n+1}\right)|\mathfrak{F}_n\right) = \mathbb{E}\left(Z_{\tau_k}|\mathfrak{F}_n\right) \leq Y_n, \qquad \mathbb{P}-\text{p.n.}$$
Czyli
$$\mathbb{E}\left(Y_{n+1}|\mathfrak{F}_n\right) \leq Y_n, \qquad \mathbb{P}-\text{p.n.}$$
Ponieważ $Y_n \geq Z_n$ więc
$$Y_n \geq \max\left\{Z_n, \mathbb{E}\left(Y_{n+1}|\mathfrak{F}_n\right)\right\}, \qquad \mathbb{P}-\text{p.n.}$$
Pozostaje do wykazania nierówność odwrotna
$$Y_n \leq \max\left\{Z_n, \mathbb{E}\left(Y_{n+1}|\mathfrak{F}_n\right)\right\}, \qquad \mathbb{P}-\text{p.n.}$$
Z definicji Y_n jako istotnego kresu górnego $\mathbb{E}\left(Z_\tau, |\mathfrak{F}_n\right)$ wystarczy pokazać, że dla dowolnego $\tau \in \Sigma^n$,
$$\mathbb{E}\left(Z_\tau, |\mathfrak{F}_n\right) \leq \max\left\{Z_n, \mathbb{E}\left(Y_{n+1}|\mathfrak{F}_n\right)\right\}, \qquad \mathbb{P}-\text{p.n.}$$
Mamy
$$Z_\tau = Z_n\chi_{\{\tau=n\}} + Z_\tau\chi_{\{\tau>n\}} = Z_n\chi_{\{\tau=n\}} + Z_{\tau\vee(n+1)}\chi_{\{\tau>n\}}, \qquad \mathbb{P}-\text{p,n.}$$
Stąd $\mathbb{P}$-p.n. zachodzi
$$\begin{aligned}\mathbb{E}\left(Z_\tau|\mathfrak{F}_n\right) &= Z_n\chi_{\{\tau=n\}} + \chi_{\{\tau>n\}}\mathbb{E}\left(Z_{\tau\vee(n+1)}|\mathfrak{F}_n\right)\\ &= Z_n\chi_{\{\tau=n\}} + \chi_{\{\tau>n\}}\mathbb{E}\left(\mathbb{E}\left(Z_{\tau\vee(n+1)}|\mathfrak{F}_{n+1}\right)|\mathfrak{F}_n\right)\\ &\leq Z_n\chi_{\{\tau=n\}} + \chi_{\{\tau>n\}}\mathbb{E}\left(Y_{n+1}|\mathfrak{F}_n\right)\\ &\leq \max\left\{Z_n, \mathbb{E}\left(Y_{n+1}|\mathfrak{F}_n\right)\right\}.\end{aligned}$$

Część (c) wynika z definicji Y_n oraz z równości (7.1).

Pokażemy teraz część (d). Z punktów (a) i (b) wynika, że $(Y_n, \mathfrak{F}_n)$ jest nadmartyngałem majoryzującym (Z_n). Załóżmy, że $\inf_n Z_n \in L^1(\Omega, \mathfrak{F}, \mathbb{P})$. Niech $(\tilde{Y}_n, \mathfrak{F}_n)$ będzie nadmartyngałem majoryzującym (Z_n). Z tego, że $\tilde{Y}_n \geq Z_n$, $\mathbb{P}$-p.n. dla wszystkich n wynika $\tilde{Y}_n \geq \inf_{k\in\mathbb{N}} Z_k$, a w konsekwencji, ponieważ $\inf_{k\in\mathbb{N}} Z_k \in L^1(\Omega, \mathfrak{F}, \mathbb{P})$,
$$\tilde{Y}_n \geq \mathbb{E}\left(\inf_{k\in\mathbb{N}} Z_k|\mathfrak{F}_n\right), \qquad \mathbb{P}-\text{p.n.},\ \forall n.$$
Czyli $(\tilde{Y}_n, \mathfrak{F}_n)$ jest nadmartyngałem prawostronnie domkniętym. Stąd, z twierdzenia Dooba o stopowaniu (patrz Twierdzenie 7.3), dla dowolnego n i momentu Markowa $\tau \geq n$,

$$\tilde{Y}_n \geq \mathbb{E}\left(\tilde{Y}_\tau|\mathfrak{F}_n\right) \geq \mathbb{E}\left(Z_\tau|\mathfrak{F}_n\right), \qquad \mathbb{P}-\text{p.n.}$$

Stąd, patrz Uwaga 7.1,

$$\tilde{Y}_n \geq \operatorname*{ess\,sup}_{\tau\in\tilde{\Sigma}^n} \mathbb{E}\left(Z_\tau|\mathfrak{F}_n\right) = \operatorname*{ess\,sup}_{\tau\in\tilde{\Sigma}^n} \mathbb{E}\left(Z_\tau|\mathfrak{F}_n\right) := Y_n, \qquad \mathbb{P}-\text{p.n.} \qquad \square$$

Do dowodu ostatniej części twierdzenia będziemy potrzebowali trzech lematów.

Lemat 7.3 *Dla dowolnego momentu stopu $\tau \leq \tau_0$, $(Y_{n\wedge\tau}, \mathfrak{F}_n)$ jest martyngałem. Ponadto, jeżeli $\tau < \infty$, $\mathbb{P}$-p.n., to $Y_\tau \in L^1(\Omega, \mathfrak{F}, \mathbb{P})$ oraz $\mathbb{E}\, Y_\tau \geq \mathbb{E}\, Y_0$.*

Dowód Oczywiście dla dowolnego n i τ zmienna losowa $Y_{n\wedge\tau}$ jest całkowalna i $\mathfrak{F}_n$-mierzalna. Mamy

$$\mathbb{E}\left(Y_{(n+1)\wedge\tau}|\mathfrak{F}_n\right) = Y_{n\wedge\tau}\chi_{\{\tau\leq n\}} + \chi_{\{\tau>n\}}\mathbb{E}\left(Y_{n+1}|\mathfrak{F}_n\right).$$

Ponieważ $\tau \leq \tau_0$, więc na zbiorze $\{\tau > n\}$ mamy $\{\tau_0 > n\}$ a więc $Y_n > Z_n$. Stąd

$$\chi_{\{\tau>n\}}Y_n = \chi_{\{\tau>n\}} \max\left\{Z_n, \mathbb{E}\left(Y_{n+1}|\mathfrak{F}_n\right)\right\} = \chi_{\{\tau>n\}}\mathbb{E}\left(Y_{n+1}|\mathfrak{F}_n\right).$$

W konsekwencji

$$\mathbb{E}\left(Y_{(n+1)\wedge\tau}|\mathfrak{F}_n\right) = Y_{n\wedge\tau}\chi_{\{\tau\leq n\}} + \chi_{\{\tau>n\}}Y_n = Y_{n\wedge\tau}.$$

Pokazaliśmy, że $(Y_{\tau\wedge n}, \mathfrak{F}_n)$ jest martyngałem. W szczególności,

$$\mathbb{E}\, Y_{n\wedge\tau} = \mathbb{E}\, Y_0, \qquad \forall\, n \in \mathbb{N}.$$

Jeżeli $\tau < \infty$, $\mathbb{P}$-p.n. to $Y_{n\wedge\tau} \to Y_\tau$, $\mathbb{P}$-p.n. Wystarczy więc uzasadnić, że

$$\lim_{n\to+\infty} \mathbb{E}\, Y_{n\wedge\tau} \leq \mathbb{E} \lim_{n\to+\infty} Y_{n\wedge\tau} = \mathbb{E}\, Y_\tau.$$

W tym celu niech

$$U_n := \mathbb{E}\left(\sup_{k\in\mathbb{N}} Z_k^+|\mathfrak{F}_n\right), \qquad n \in \mathbb{N}.$$

Z definicji Y_n jest istotnym kresem górnym $\mathbb{E}\left(Z_\sigma|\mathfrak{F}_n\right)$, $\sigma \in \Sigma^n$. Stąd

$$Y_n \leq U_n, \qquad \mathbb{P}-\text{p.n.}$$

Tak więc $U_n - Y_n \geq 0$ i stosując lemat Fatou otrzymujemy

$$\liminf_{n\to\infty} \mathbb{E}\left(U_{n\wedge\tau} - Y_{n\wedge\tau}\right) \geq \mathbb{E}\left(U_\tau - Y_\tau\right).$$

Ponieważ $(U_n, \mathfrak{F}_n)$ jest martyngałem prawostronnie domkniętym $U_\tau \in L^1(\Omega, \mathfrak{F}, \mathbb{P})$ oraz
$$\mathbb{E}\, U_{\tau\wedge n} = \mathbb{E}\, U_\tau.$$
Z martyngałowości $(Y_{n\wedge\tau})$ wynika, że $\mathbb{E}\, Y_{n\wedge\tau} = \mathbb{E}\, Y_0$. Stąd
$$\liminf_{n\to\infty} \mathbb{E}\,(U_{n\wedge\tau} - Y_{n\wedge\tau}) = \mathbb{E}\, U_\tau - \mathbb{E}\, Y_0 \geq \mathbb{E}\,(U_\tau - Y_\tau) = \mathbb{E}\, U_\tau - \mathbb{E}\, Y_\tau,$$
przy czym $\mathbb{E}\, Y_\tau$ jest dobrze określone bo $Y_\tau \leq U_\tau$ i U_τ jest zmienną losową całkowalną.

Podsumowując mamy
$$\mathbb{E}\, U_\tau - \mathbb{E}\, Y_0 \geq \mathbb{E}\, U_\tau - \mathbb{E}\, Y_\tau \quad \text{oraz} \quad U_\tau \geq Y_\tau, \quad \text{oraz} \quad U_\tau \in L^1(\Omega, \mathfrak{F}, \mathbb{P}).$$
Stąd
$$\mathbb{E}\, Y_0 \leq \mathbb{E}\, Y_\tau \leq \mathbb{E}\, U_\tau.$$
Ponieważ $\mathbb{E}\,|Y_0| < \infty$, więc $Y_\tau \in L^1(\Omega, \mathfrak{F}, \mathbb{P})$, co kończy dowód lematu. □

Lemat 7.4 *Jeżeli τ jest momentem Markowa względem filtracji $(\mathfrak{F}_n)$ to dla dowolnej całkowalnej zmiennej losowej ξ zachodzi*
$$\mathbb{E}\,(\xi|\mathfrak{F}_\tau) = \sum_{n=0}^{\infty} \chi_{\{\tau=n\}} \mathbb{E}\,(\xi|\mathfrak{F}_n) + \chi_{\{\tau=\infty\}}\xi, \qquad \mathbb{P} - p.n.$$

Dowód Oznaczmy przez η prawą stronę powyższej równości. Mamy
$$\{\eta \leq x\} \cap \{\tau = k\} = \{\mathbb{E}\,(\xi|\mathfrak{F}_k) \leq x\} \cap \{\tau = k\} \in \mathfrak{F}_k, \qquad \forall\, x \in \mathbb{R},\ \forall\, k \in \mathbb{N}.$$
Stąd η jest $\mathfrak{F}_\tau$-mierzalna.

Bez straty ogólności możemy założyć, że $\xi \geq 0$. Niech $A \in \mathfrak{F}_\tau$. Ponieważ $A \cap \{\tau = k\} \in \mathfrak{F}_k$, dla dowolnego k, zachodzi
$$\begin{aligned}\int_A \eta\, d\mathbb{P} &= \sum_{k=0}^{\infty} \int_{A\cap\{\tau=k\}} \mathbb{E}\,(\xi|\mathfrak{F}_k)\, d\mathbb{P} + \int_{A\cap\{\tau=\infty\}} \xi d\mathbb{P} \\ &= \sum_{k=0}^{\infty} \int_{A\cap\{\tau=k\}} \xi d\mathbb{P} + \int_{A\cap\{\tau=\infty\}} \xi d\mathbb{P} \\ &= \int_A \xi d\mathbb{P}. \qquad \square\end{aligned}$$

Lemat 7.5 *Dla dowolnego $\tau \in \Sigma$,*
$$Y_\tau = \operatorname*{ess\,sup}_{\sigma\in\Sigma^\tau} \mathbb{E}\,(Z_\sigma|\mathfrak{F}_\tau)\,, \qquad \mathbb{P} - p.n.$$
oraz
$$\mathbb{E}\, Y_\tau = \sup_{\sigma\in\Sigma^\tau} \mathbb{E}\, Z_\sigma.$$

Dowód Niech $\sigma \in \Sigma^\tau$. Z poprzedniego lematu wynika, że dla dowolnego n,

$$\chi_{\{\tau=n\}}\mathbb{E}\left(Z_\sigma|\mathfrak{F}_\tau\right) = \chi_{\{\tau=n\}}\mathbb{E}\left(Z_\sigma|\mathfrak{F}_n\right), \qquad \mathbb{P}-\text{p.n.}$$

Następnie

$$\begin{aligned}\chi_{\{\tau=n\}}\mathbb{E}\left(Z_\sigma|\mathfrak{F}_\tau\right) &= \mathbb{E}\left(\chi_{\{\tau=n\}}Z_\sigma|\mathfrak{F}_\tau\right) = \mathbb{E}\left(\chi_{\{\tau=n\}}Z_{\sigma\vee n}|\mathfrak{F}_\tau\right) = \chi_{\{\tau=n\}}\mathbb{E}\left(Z_{\sigma\vee n}|\mathfrak{F}_\tau\right)\\ &= \chi_{\{\tau=n\}}\mathbb{E}\left(Z_{\sigma\vee n}|\mathfrak{F}_n\right) \le \chi_{\{\tau=n\}}Y_n = \chi_{\{\tau=n\}}Y_\tau.\end{aligned}$$

Stąd, ponieważ nierówność zachodzi dla każdego $\sigma \in \Sigma^\tau$ i dla każdego n mamy

$$\operatorname*{ess\,sup}_{\sigma\in\Sigma^\tau}\mathbb{E}\left(Z_\sigma|\mathfrak{F}_\tau\right) \le Y_\tau.$$

Stąd

$$\sup_{\sigma\in\Sigma^\tau}\mathbb{E}\,Z_\sigma \le Y_\tau, \qquad \mathbb{P}-\text{p.n.}$$

Pokazujemy nierówność przeciwną. Mamy

$$\chi_{\{\tau=n\}}Y_\tau = \chi_{\{\tau=n\}}Y_n = \chi_{\{\tau=n\}}\operatorname*{ess\,sup}_{\sigma\in\Sigma^n}\mathbb{E}\left(Z_\sigma|\mathfrak{F}_n\right), \qquad \mathbb{P}-\text{p.n.}$$

Stąd wystarczy pokazać, że dla dowolnych $n \in \mathbb{N}$ i $\sigma \in \Sigma^n$ zachodzi

$$\chi_{\{\tau=n\}}\mathbb{E}\left(Z_\sigma|\mathfrak{F}_n\right) \le \chi_{\{\tau=n\}}\operatorname*{ess\,sup}_{\tilde\sigma\in\Sigma^\tau}\mathbb{E}\left(Z_{\tilde\sigma}|\mathfrak{F}_\tau\right).$$

Mamy

$$\begin{aligned}\chi_{\{\tau=n\}}\mathbb{E}\left(Z_\sigma|\mathfrak{F}_n\right) &= \mathbb{E}\left(Z_\sigma\chi_{\{\tau=n\}}|\mathfrak{F}_n\right) = \mathbb{E}\left(Z_{\sigma\vee\tau}\chi_{\{\tau=n\}}|\mathfrak{F}_n\right)\\ &\le \operatorname*{ess\,sup}_{\tilde\sigma\in\Sigma^\tau}\mathbb{E}\left(Z_{\tilde\sigma}\chi_{\{\tau=n\}}|\mathfrak{F}_n\right) = \operatorname*{ess\,sup}_{\tilde\sigma\in\Sigma^\tau}\chi_{\{\tau=n\}}\mathbb{E}\left(Z_{\tilde\sigma}|\mathfrak{F}_n\right)\\ &\le \chi_{\{\tau=n\}}\operatorname*{ess\,sup}_{\tilde\sigma\in\Sigma^\tau}\mathbb{E}\left(Z_{\tilde\sigma}|\mathfrak{F}_\tau\right). \qquad \square\end{aligned}$$

Dowód części (e) Niech $\widehat{\tau}$ będzie optymalną regułą zatrzymania. Pokażemy, że $\widehat{\tau} \ge \tau_0$, $\mathbb{P}$-p.n. Z Lematu 7.5 mamy

$$\mathbb{E}\,Y_{\widehat{\tau}} = \sup_{\sigma\in\Sigma^{\widehat{\tau}}}\mathbb{E}\,Z_\sigma \le \mathbb{E}\,Z_{\widehat{\tau}}.$$

Z drugiej strony $\widehat{\tau} \in \Sigma^{\widehat{\tau}}$. Stąd

$$\mathbb{E}\,Y_{\widehat{\tau}} = \mathbb{E}\,Z_{\widehat{\tau}}.$$

Ponieważ $Y_{\widehat{\tau}} \ge Z_{\widehat{\tau}}$, $\mathbb{P}$-p.n., więc $Y_{\widehat{\tau}} = Z_{\widehat{\tau}}$, $\mathbb{P}$-p.n.. Pozostaje do pokazania, że jeżeli $\tau_0 < \infty$, $\mathbb{P}$-p.n. to τ_0 jest optymalną regułą zatrzymania, to znaczy, że

$$\mathbb{E}\,Z_{\tau_0} = \sup_{\tau\in\Sigma}\mathbb{E}\,Z_\tau.$$

Z definicji τ_0 i z tego, że $\tau_0 < \infty$, $\mathbb{P}$-p.n. wnioskujemy, że

$$Y_{\tau_0} = Z_{\tau_0}, \qquad \mathbb{P} - \text{p.n.}$$

a więc w szczególności

$$\mathbb{E}\, Y_{\tau_0} = \mathbb{E}\, Z_{\tau_0}. \tag{7.5}$$

Z Lematu 7.3,

$$\mathbb{E}\, Y_{\tau_0} \geq \mathbb{E}\, Y_0. \tag{7.6}$$

Z definicji obwiedni Snella,

$$Y_0 = \operatorname*{ess\,sup}_{\tilde{\tau}\in\Sigma^0=\Sigma} \mathbb{E}\left(Z_\sigma|\mathfrak{F}_0\right).$$

Więc, patrz Twierdzenie 7.1,

$$\mathbb{E}\, Y_0 = \sup_{\sigma\in\Sigma} \mathbb{E}\,\mathbb{E}\left(Z_\sigma|\mathfrak{F}_0\right) = \sup_{\sigma\in\Sigma} \mathbb{E}\, Z_\sigma.$$

W konsekwencji, z (7.5) i (7.6) wynika, że

$$\mathbb{E}\, Z_{\tau_0} \geq \sup_{\sigma\in\Sigma} \mathbb{E}\, Z_\sigma,$$

czyli τ_0 jest optymalną regułą zatrzymania. □

7.1 Nadmartyngały prawostronnie domknięte

Definicja 7.2 Nadmartyngał $(Y_n, \mathfrak{F}_n)$ jest *prawostronnie domknięty*, jeżeli istnieje zmienna losowa $Y_\infty \in L^1(\Omega, \mathfrak{F}, \mathbb{P})$, taka że

$$Y_n \geq \mathbb{E}\left(Y_\infty, |\mathfrak{F}_n\right), \qquad \mathbb{P} - \text{p.n.}$$

Mowimy wtedy, że $(Y_n, \mathfrak{F}_n)$ jest *prawostronnie domknięty przez* Y_∞.

Martyngał $(Y_n, \mathfrak{F}_n)$ jest *prawostronnie domknięty*, jeżeli istnieje zmienna losowa $Y_\infty \in L^1(\Omega, \mathfrak{F}, \mathbb{P})$, taka że

$$Y_n = \mathbb{E}\left(Y_\infty, |\mathfrak{F}_n\right), \qquad \mathbb{P} - \text{p.n.}$$

Mówimy wtedy, że $(Y_n, \mathfrak{F}_n)$ jest *prawostronnie domknięty przez* Y_∞.

Dla momentu Markowa τ i prawostronnie domkniętego nadmartyngału $(Y_n, \mathfrak{F}_n)$ domkniętego przez Y_∞, kładziemy $Y_\tau = Y_\infty$ na zbiorze $\{\tau = \infty\}$.

Twierdzenie 7.3 *Niech $\tau_1 \leq \tau_2$ będą momentami Markowa.*
(a) Jeśli $(Y_n, \mathfrak{F}_n)$ jest martyngałem prawostronnie domknięty przez Y_∞, to zmienne losowe Y_{τ_i}, $i = 1, 2$, są całkowalne oraz

$$Y_{\tau_1} = \mathbb{E}\left(Y_{\tau_2}|\mathfrak{F}_{\tau_1}\right) = \mathbb{E}\left(Y_\infty|\mathfrak{F}_{\tau_1}\right), \qquad \mathbb{P} - p.n.$$

(b) Jeśli $(Y_n, \mathfrak{F}_n)$ jest nadmartyngałem prawostronnie domknięty przez Y_∞, to zmienne losowe Y_{τ_i}, $i = 1, 2$, są całkowalne oraz

$$Y_{\tau_1} \geq \mathbb{E}\left(Y_{\tau_2}|\mathfrak{F}_{\tau_1}\right), \qquad \mathbb{P} - p.n.$$

Dowód (a)

7.2 Ciekawy przykład

Niech (ξ_n) będzie ciągiem niezależnych zmiennych losowych o tym samym rozkładzie μ. Niech $c > 0$ oraz niech

$$Z_n := \max_{k\le n} \xi_k - cn, \qquad n \in \mathbb{N}.$$

7.3 Wycena opcji amerykańskich

Niech $S(n)$, $n = 0, 1, \ldots$, oznacza (losową) cenę instrumentu bazowego w chwili n. Niech $r(n)$ oznacza (być może losową i zależną od czasu n) stopę krótką na odcinek od n do $n+1$. Przypomnijmy, że $B(n)$ to wartość rachunku bankowego w chwili n przy założeniu, że $B(0) = 1$. Czyli

$$B(n) = \exp\left\{\delta \sum_{k=0}^{n-1} r(k)\right\}, \qquad n = 0, 1, \ldots,$$

gdzie δ oznacza długość okresu rozliczeniowego. Mamy

$$S(n-1) = \exp\{-\delta r(n-1)\}\, \mathbb{E}^Q\left(S(n)|\mathfrak{F}_{n-1}\right), \qquad n = 1, \ldots,$$

gdzie Q oznacza miarę wolną od ryzyka (inaczej miarę martyngałową).

Rozważmy amerykańską opcję sprzedaży (opcja put) z momentem wygaśnięcia opcji N i ceną realizacji K. Niech

$$I(n) = (K - S(n))^+, \qquad n = 0, \ldots, N.$$

Niech $P(n)$ oznacza cenę opcji w chwili $n \le N$ przy założeniu, że opcja jest aktywna w chwili n. Oczywiście

$$P(N) = I(N).$$

Ile wynosi $P(N-1)$? W chwili $N-1$ posiadacz opcji może ją zrealizować otrzymując $I(N-1)$ lub odłożyć decyzje na później. Inaczej ma prawo albo do $I(n-1)$ albo do opcji europejskiej o cenie realizacji $P(N)$ w momencie wygaśnięcia N. Stąd wartość opcji w chwili $N-1$ wynosi

$$P(N-1) = \max\left\{I(N-1), \exp\{-\delta r(N-1)\}\, \mathbb{E}^Q\left(P(N)|\mathfrak{F}_{N-1}\right)\right\}.$$

Ogólnie

$$P(n) = \max\left\{I(n), \exp\{-\delta r(n)\}\, \mathbb{E}^Q\left(P(n+1)|\mathfrak{F}_n\right)\right\}, \qquad n = 0, 1, \ldots, N-1.$$

Niech

$$\tilde{P}(n) = \frac{P(n)}{B(n)}, \qquad \tilde{I}(n) = \frac{I(n)}{B(n)}, \qquad n = 0, 1, \ldots, N.$$

oznaczają zdyskontowane do czasu 0 procesy cen.

Wówczas $\tilde{P}(0) = P(0)$, oraz

$$\begin{aligned}
&\tilde{P}(N) = \tilde{I}(N),\\
&\tilde{P}(n) = \max\left\{\tilde{I}(n), \mathbb{E}^Q\left(\tilde{P}(n+1)|\mathfrak{F}_n\right)\right\}, \qquad n = 0, 1, \ldots, N-1.
\end{aligned}$$

Stąd $\tilde{P}$ jest obwiednią Snella $\tilde{I}$.

Zadanie 7.2 Rozważmy model CRR cen akcji. Niech $S(0) = 100$, $R = 0.06$, $U = 15\%$ a $D = -10\%$. Rozważmy amerykańską opcję sprzedaży o czasie wygaśnięcia $N = 3$ i cene wykupu $K = 100$. Znaleźć $P(0)$.

Odpowiedź Przypomnijmy, że w modelu CRR stopa R jest prosta (to znaczy bez ciągłej kapitalizacji odsetek) i stała w czasie. Stąd rachunek bankowy dany jest wzorem

$$B(n) = (1+R)^n\,, \qquad n \in \mathbb{N}.$$

Przypomnijmy, że w modelu CRR drzewo jest rekombinowane,

$$S(n+1; \mathrm{w}_n\mathrm{u}) = S(n; \mathrm{w}_n)\,(1+U)\,, \qquad S(n+1; \mathrm{w}_n\mathrm{d}) = S(n; \mathrm{w}_n)\,(1+D)$$

oraz, że w modelu CRR prawdopodobieństwo martyngałowe $p(n; w_n)$ jest stałe i równe

$$p = \frac{R-D}{U-D} = \frac{0,06+0.1}{0,15+0,1} = \frac{0,16}{0.25} = 0,64.$$

Liczymy drzewo cen $S(n)$, wypłat $I(n) = (K - S(n))^+$ oraz zdyskontowanych wypłat $\tilde{I}(n) = I(n)\,(1+R)^{-n}$. Mamy

$$\begin{array}{ccccccc}
 & & & & & & \begin{array}{l} S(3;\mathrm{uuu}) = 152,09\\ I(3;\mathrm{uuu}) = 0\\ \tilde{I}(3;\mathrm{uuu}) = 0\end{array}\\
 & & & & \begin{array}{l} S(2;\mathrm{uu}) = 132,24\\ I(2;\mathrm{uu}) = 0\\ \tilde{I}(2;\mathrm{uu}) = 0\end{array} & \nearrow & \\
 & & \begin{array}{l} S(1;\mathrm{u}) = 115\\ I(1;\mathrm{u}) = 0\\ \tilde{I}(1;\mathrm{u}) = 0\end{array} & \nearrow & & \searrow & \\
 & \nearrow & & & & & \begin{array}{l} S(3;\mathrm{uud}) = 119,025\\ I(3;\mathrm{uud}) = 0\\ \tilde{I}(3;\mathrm{uud}) = 0\end{array}\\
\begin{array}{l} S(0) = 100\\ I(0) = 0\\ \tilde{I}(0) = 0\end{array} & & & \searrow & & \nearrow & \\
 & & & & \begin{array}{l} \S(2;\mathrm{ud}) = 103,5\\ I(2;\mathrm{ud}) = 0\\ \tilde{I}(2;\mathrm{ud}) = 0\end{array} & \searrow & \\
 & \searrow & & \nearrow & & & \begin{array}{l} S(3;\mathrm{udd}) = 93,15\\ I(3;\mathrm{udd}) = 6,85\\ \tilde{I}(3;\mathrm{udd}) = 5,75\end{array}\\
 & & \begin{array}{l} S(1;\mathrm{d}) = 90\\ I(1;\mathrm{d}) = 10\\ \tilde{I}(1;\mathrm{d}) = 9,43\end{array} & \searrow & & \nearrow & \\
 & & & & \begin{array}{l} S(2;\mathrm{dd}) = 81\\ I(2;\mathrm{dd}) = 19\\ \tilde{I}(2;\mathrm{dd}) = 16,91\end{array} & \searrow & \\
 & & & & & & \begin{array}{l} S(3;\mathrm{ddd}) = 72,9\\ I(3;\mathrm{ddd}) = 27,1\\ \tilde{I}(3;\mathrm{ddd}) = 22,75\end{array}
\end{array}$$

Oznaczmy przez

$$E(n) = \mathbb{E}^Q\left(\tilde{P}(n+1)|\mathfrak{F}_n\right).$$

warunkową wartość oczekiwaną zdyskontowanej ceny opcji w chwili $n+1$ ze względu na miarę martyngałową. Drzewo zdyskontowanych cen $\tilde{P}$ opcji amerykańskiej, która jest obwiednią Snella procesu $\tilde{I}$ wyliczamy następująco:

$$\begin{aligned}
\tilde{P}(3) &= \tilde{I}(3), \\
\tilde{P}(n;\mathrm{w}_n) &= \max\left\{\tilde{I}(n;\mathrm{w}_n), E(n;\mathrm{w}_n)\right\}, \quad n=0,1,2, \\
E(n;\mathrm{w}_n) &= p\tilde{P}(n+1;\mathrm{w}_n\mathrm{u}) + (1-p)\tilde{P}(n+1;\mathrm{w}_n\mathrm{d}) \quad n=0,1,2.
\end{aligned}$$

W efekcie otrzymujemy

$$\begin{array}{ccccccc}
 & & & & & \nearrow & \tilde{P}(3;\mathrm{uuu}) = \mathbf{0} \\
 & & & & \begin{array}{l} E(2;\mathrm{uu}) = 0 \\ \tilde{P}(2;\mathrm{uu}) = 0 \end{array} & & \\
 & & \begin{array}{l} E(1;\mathrm{u}) = 0,75 \\ \tilde{P}(1;\mathrm{u}) = 0,75 \end{array} & \begin{array}{l}\nearrow \\ \searrow\end{array} & & \searrow & \\
 & \nearrow & & & & \nearrow & \tilde{P}(3;\mathrm{uud}) = \mathbf{0} \\
\begin{array}{l} E(0) = 3,87 \\ \tilde{P}(0) = 3,87 \end{array} & & & & \begin{array}{l} E(2;\mathrm{ud}) = 2,07 \\ \tilde{P}(2;\mathrm{ud}) = 2,07 \end{array} & \searrow & \\
 & \searrow & & \nearrow & & & \tilde{P}(3;\mathrm{udd}) = \mathbf{5,75} \\
 & & \begin{array}{l} E(1;\mathrm{d}) = 7,42 \\ \tilde{P}(1;\mathrm{d}) = \mathbf{9,43} \end{array} & \searrow & \begin{array}{l} E(2;\mathrm{dd}) = 11,87 \\ \tilde{P}(2;\mathrm{dd}) = \mathbf{16,91} \end{array} & \nearrow & \\
 & & & & & \searrow & \tilde{P}(3;\mathrm{ddd}) = \mathbf{22,75},
\end{array}$$

gdzie liczby pogrubione występują w chwili gdy $\tilde{P}(n) = \tilde{I}(n)$.

Zadanie 7.3 Niech $B(t,T)$ oznacza cenę w chwili t obligacji zerokuponowej o momencie zapadalności $T > t$.[1] Załóżmy, że długość okresu rozliczeniowego wynosi 1, a drzewo cen obligacji jest następujące

$$\begin{array}{ccccc}
 & & & & \begin{array}{l} B(2,2;\mathrm{uu}) = 1 \\ B(2,3;\mathrm{uu}) = 0,989 \end{array} \\
 & & \begin{array}{l} B(1,1;\mathrm{u}) = 1 \\ B(1,2;\mathrm{u}) = 0,9983 \\ B(1,3;\mathrm{u}) = 0,908 \end{array} & \begin{array}{c} \nearrow \\ r(1;\mathrm{u}) = 2\% \\ \searrow \end{array} & \\
 & \nearrow & & & \begin{array}{l} B(2,2;\mathrm{ud}) = 1 \\ B(2,3;\mathrm{ud}) = 0,89 \end{array} \\
\begin{array}{l} B(0,1) = 0,9967 \\ B(0,2) = 0,9940 \\ B(0,3) = 0,87 \end{array} & r(0) = 4\% & & & \\
 & \searrow & & & \begin{array}{l} B(2,2;\mathrm{du}) = 1 \\ B(2,3;\mathrm{du}) = 0,985 \end{array} \\
 & & \begin{array}{l} B(1,1;\mathrm{d}) = 1 \\ B(1,2;\mathrm{d}) = 0,9967 \\ B(1,3;\mathrm{d}) = 0,85 \end{array} & \begin{array}{c} \nearrow \\ r(1;\mathrm{d}) = 4\% \\ \searrow \end{array} & \\
 & & & & \begin{array}{l} B(2,2;\mathrm{dd}) = 1 \\ B(2,3;\mathrm{dd}) = 0,85. \end{array}
\end{array} .$$

Znaleźć drzewo cen opcji amerykańskiej put na oblikację o czasie zapadalności 3, czasie wygaśnięcia opcji 2 i cenie sprzedaży $K = 0,92$.

Odpowiedź Powinniśmy najpierw sprawdzić czy w modelu nie ma arbitrażu. W tym celu należy pokazać, że prawdopodobieństwa (martyngałowe) $p(t;w_t)$ wyliczone ze wzoru

$$\begin{aligned}
&B(t,T;\mathrm{w}_t) \\
&\quad = B(t,t+1;\mathrm{w}_t)\left(p(t;\mathrm{w}_t)B(t+1;T;\mathrm{w}_t\mathrm{u}) + (1-p(t;\mathrm{w}_t))\,B(t+1;T;\mathrm{w}_t\mathrm{d})\right),
\end{aligned}$$

[1] Obligacja ta daję posiadaczowi 1 w chwili T.

nie zależy od $T = t+2, t+3, \ldots$ i jest z przedziału $(0, 1)$.

Mamy

$$
\begin{array}{lllll}
 & & & p(1;\mathrm{u}) = 0,1971 & B(2,2;\mathrm{uu}) = 1 \\
 & & B(1,1;\mathrm{u}) = 1 & \nearrow & B(2,3;\mathrm{uu}) = 0,989 \\
 & p(0) = 0,3949 & B(1,2;\mathrm{u}) = 0,9983 & r(1;\mathrm{u}) = 2\% & \\
 & \nearrow & B(1,3;\mathrm{u}) = 0,908 & \searrow & B(2,2;\mathrm{ud}) = 1 \\
B(0,1) = 0,9967 & & & q(1;\mathrm{u}) = 0,8029 & B(2,3;\mathrm{ud}) = 0,89 \\
B(0,2) = 0,9940 & r(0) = 4\% & & & \\
B(0,3) = 0,87 & & & p(1;\mathrm{d}) = 0,0210 & B(2,2;\mathrm{du}) = 1 \\
 & \searrow & B(1,1;\mathrm{d}) = 1 & \nearrow & B(2,3;\mathrm{du}) = 0,985 \\
 & q(0) = 0,6051 & B(1,2;\mathrm{d}) = 0,9967 & r(1;\mathrm{d}) = 4\% & \\
 & & B(1,3;\mathrm{d}) = 0,85 & \searrow & B(2,2;\mathrm{dd}) = 1 \\
 & & & q(1;\mathrm{d}) = 0,9790 & B(2,3;\mathrm{dd}) = 0,85.
\end{array}
$$

Następnie wypisujemy drzewo wypłat

$$I(t;\mathrm{w}_t) = (0,98 - B(t,3;\mathrm{w}_t)^+$$

drzewo wypłat zdyskontowanych

$$\tilde{I}(t;\mathrm{w}_t) = B(0,1)B(1,2,\mathrm{w}_1) \times \ldots \times B(t-1, t\mathrm{w}_{t-1})I(t;\mathrm{w}_t);$$

$$
\begin{array}{lllll}
 & & & p(1;\mathrm{u}) = 0,1971 & I(2;\mathrm{uu}) = 0 \\
 & & I(1,;\mathrm{u}) = 0,072 & \nearrow & \tilde{I}(2;\mathrm{uu}) = 0 \\
 & p(0) = 0,3949 & \tilde{I}(1;\mathrm{u}) = 0,0717 & r(1;\mathrm{u}) = 2\% & \\
 & \nearrow & B(1,3;\mathrm{u}) = 0,908 & \searrow & I(2;\mathrm{ud}) = 0,09 \\
I(0) = 0,11 & & & q(1;\mathrm{u}) = 0,8029 & \tilde{I}(2;\mathrm{ud}) = 0,089 \\
\tilde{I}(0) = 0,11 & r(0) = 4\% & & & \\
B(0,3) = 0,87 & & & p(1;\mathrm{d}) = 0,0210 & I(2;\mathrm{du}) = 0 \\
 & \searrow & I(1;\mathrm{d}) = 0,13 & \nearrow & \tilde{I}(2;\mathrm{du}) = 0 \\
 & q(0) = 0,6051 & \tilde{I}(1;\mathrm{d}) = 0,1296 & r(1;\mathrm{d}) = 4\% & \\
 & & B(1,3;\mathrm{d}) = 0,85 & \searrow & I(2;\mathrm{dd}) = 0,13 \\
 & & & q(1;\mathrm{d}) = 0,9790 & \tilde{I}(2;\mathrm{dd}) = 0,129.
\end{array}
$$

W końcu otrzymujemy

$$
\begin{array}{lllll}
 & & & p(1;\mathrm{u}) = 0,1971 & \tilde{I}(2;\mathrm{uu}) = 0 \\
 & & E(1,;\mathrm{u}) = 0,0714 & \nearrow & \tilde{P}(2;\mathrm{uu}) = \mathbf{0} \\
 & p(0) = 0,3949 & \tilde{I}(1;\mathrm{u}) = 0,0717 & r(1;\mathrm{u}) = 2\% & \\
 & \nearrow & \tilde{P}(1;\mathrm{u}) = \mathbf{0,0717} & \searrow & \tilde{I}(2;\mathrm{ud}) = 0,09 \\
E(0) = 0,1067 & & & q(1;\mathrm{u}) = 0,8029 & \tilde{P}(2;\mathrm{ud}) = \mathbf{0,09} \\
\tilde{I}(0) = 0,11 & r(0) = 4\% & & & \\
\tilde{P}(0) = \mathbf{0,11} & & & p(1;\mathrm{d}) = 0,0210 & \tilde{I}(2;\mathrm{du}) = 0 \\
 & \searrow & E(1;\mathrm{d}) = 0,1263 & \nearrow & \tilde{P}(2;\mathrm{du}) = \mathbf{0} \\
 & q(0) = 0,6051 & \tilde{I}(1;\mathrm{d}) = 0,1296 & r(1;\mathrm{d}) = 4\% & \\
 & & \tilde{P}(1;\mathrm{d}) = \mathbf{0{,}1296} & \searrow & \tilde{I}(2;\mathrm{dd}) = 0,128 \\
 & & & q(1;\mathrm{d}) = 0,9790 & \tilde{P}(2;\mathrm{dd}) = 0,\mathbf{129}.
\end{array}
,
$$

gdzie

$$E(t;\mathrm{w}_t) = p(t;\mathrm{w}_t)\tilde{I}(t+1;\mathrm{w}_t\mathrm{u}) + (1 - p(t;\mathrm{w}_t))\,\tilde{I}(t+1;\mathrm{w}_t\mathrm{d}).$$

7.4 Obwiednia Snella i programowanie dynamiczne (skończony horyzont czasowy)

Rozważmy problem stopowania łańcucha Markowa (X_n) ze skończonym horyzontem czasowym N i funkcjonałem zysku

$$J(X_0,\tau) = \mathbb{E}\left\{\sum_{k=1}^{\tau-1}\gamma^k q(X_n) + \gamma^\tau r(X_\tau)\right\}, \qquad \tau \in \Sigma_N,$$

gdzie Σ_N oznacza klasę momentów stopu $\tau \le N$. Przypomnijmy (patrz Twierdzenie **??**), że

$$\sup_{\tau\in\Sigma_N} J(X_0,\tau) = \mathbb{E}\, Q^N r(X_0),$$

gdzie

$$Qh(x) = \max\{q(x) + \gamma Ph(x), r(x)\}, \qquad x \in E.$$

Ponadto optymalny moment zatrzymania

$$\widehat{\tau} = \inf\left\{n \le N\colon Q^{N-n}r(X_n) = r(X_n)\right\}, \qquad \inf\emptyset = N.$$

Z drugiej strony

$$J(X_0,\tau) = \mathbb{E}\, Z_\tau,$$

gdzie

$$Z_n = \sum_{k=0}^{n-1}\gamma^n q(X_k) + \gamma^n r(X_n), \qquad n \le N$$

oraz $Z_n = Z_N$ dla $n \ge N$.

Tak więc założenia twierdzenia Snella są spełnione gdy

$$\mathbb{E}\,|q(X_k)| + \mathbb{E}\,|r(X_k)| < \infty, \qquad k = 0,\ldots,N.$$

Na obwiednie Snella (Y_n) mamy równania

$$Y_n = Z_n, \qquad n \ge N,$$

$$Y_n = \max\{Z_n, \mathbb{E}\,(Y_{n+1}|\mathfrak{F}_n)\}, \qquad n < N.$$

Ponieważ

$$\begin{aligned}
\mathbb{E}\,(Y_N|\mathfrak{F}_{N-1}) &= \sum_{k=0}^{N-1}\gamma^k q(X_k) + \gamma^N \mathbb{E}\,(r(X_N)|\mathfrak{F}_{N-1}) \\
&= \sum_{k=0}^{N-1}\gamma^k q(X_k) + \gamma^N Pr(X_{N-1}) \\
&= \sum_{k=0}^{N-2}\gamma^k q(X_k) + \gamma^{N-1}\left(q(X_k) + \gamma Pr(X_{N-1})\right),
\end{aligned}$$

więc

$$Y_{N-1} = \sum_{k=0}^{N-2}\gamma^k q(X_k) + Qr(X_{N-1}).$$

W konsekwencji

$$Y_n = \sum_{k=0}^{n-1}\gamma^k q(X_k) + Q^{N-n}r(X_n), \qquad n = 0,1,\ldots,N.$$

W ten sposób otrzymujemy równość $\tau_0 = \widehat{\tau}$. Oczywiście $\mathbb{P}\{\tau_0 \le N\} = 1$.

8

Twierdzenie Brussa

Niech $(\Omega, \mathfrak{F}, \mathbb{P})$ będzie przestreznią probabilistyczna oraz niech $A_1, A_2, \ldots, A_N \in \mathfrak{F}$ będą niezależnymi zdarzeniami. Niech $(\mathfrak{F}_i)$ oznacza filtrację generowaną przez zmienne losowe (χ_{A_i}). Niech Σ_N oznacza przestrzeń momentów Markowa τ względem filtracji $(\mathfrak{F}_i)$, takich, że $\mathbb{P}\left(\tau \leq N\right) = 1$. Przyjmijmy

$$J_N(\tau) := \mathbb{P}\left\{\chi_{A_\tau} = 1, \chi_{A_{\tau+1}} = 0, \ldots, \chi_{A_N} = 0\right\}.$$

Oczywiście, $J_N(\tau)$ jest prawdopodobieństwem, że τ jest ostatnią jedynką (sukcesem) w ciągu (χ_{A_i}).

Niech

$$p_i := \mathbb{P}(A_i) \neq 1, \qquad q_i := 1 - p_i, \qquad r_i = \frac{p_i}{q_i}, \qquad i = 1, \ldots, N.$$

Niech

$$t^* := \max\left\{1, \max\left\{k = 1, 2, \ldots, N \colon \sum_{j=k}^{N} r_j \geq 1\right\}\right\}$$

W poniższym twierdzeniu przyjmujemy $\inf \emptyset = 1$.

Twierdzenie 8.1 *Niech*

$$\widehat{\tau} := \min\left\{j \leq N \colon \chi_{A_j} = 1, \qquad j \geq t^*\right\}.$$

Wówczas dla każdego $\tau \in \Sigma_N$, $J_N(\tau) \leq J_N(\widehat{\tau})$ *oraz*

$$J_N(\widehat{\tau}) = \prod_{j=t^*}^{N} q_j \sum_{k=t^*}^{N} r_k.$$

Niech $(\Omega, \mathfrak{F}, \mathbb{P})$ b"edzie przestrzeni"a probabilistyczna oraz niech $A_1, A_2, \ldots, A_N$, $N < \infty$, b"edzie skończonym ci"agiem niezależnych zdarzeń z $\mathfrak{F}$.

Niech $E = \{1, \ldots, N\} \cup +\infty$. Definiujemy proces o wartościach w E wzorem

$$X_{n+1}(\omega) = \inf\left\{k > X_n(\omega)\colon \omega \in A_k\right\}.$$

Przyjmujemy, że $\inf \emptyset = +\infty$.

Niech

$$p_k = \mathbb{P}(A_k), \quad q_k = 1 - p_k = \mathbb{P}(A_k^c), \qquad r_k = \frac{p_k}{q_k}.$$

Zauważmy, że (X_n) jest jednorodnym łańcuchem Markowa na E z prawdopodobieństwem przejścia

$$P(k,l) := \mathbb{P}\left(X_{n+1} = l | X_n = k\right) = \begin{cases} p_l \prod_{j=k+1}^{l-1} q_j & \text{gdy } l < +\infty, \\ \prod_{j=k+1}^{N} q_j & \text{gdy } l = +\infty,\ k < N, \\ 1 & \text{gdy } l = +\infty,\ k \in \{N, +\infty\}. \end{cases}$$

Niech

$$Ph(k) = \begin{cases} \sum_{j=k+1}^{N} h(j)P(k,j) & \text{gdy } k \in E \setminus \{N, +\infty\}, \\ h(+\infty) & \text{gdy } k \in \{N, +\infty\} \end{cases}$$

będzie operatorem przejścia.

Niech

$$Qh(k) = \max\{Ph(k), r(k)\}, \qquad k \in E,$$

gdzie

$$r(k) = P(k, +\infty)\chi_{\{1,\ldots,N\}}(k) = \begin{cases} \prod_{j=k+1}^{N} q_j & \text{gdy } k < N, \\ 1 & \text{gdy } k = N, \\ 0 & \text{gdy } k = +\infty. \end{cases}$$

Dla nieujemnej funkcji h spełniającej $h(+\infty) = 0$ mamy

$$Qh(+\infty) = \max\{h(+\infty), 0\} = 0, \qquad Qh(N) = \max\{h(+\infty), 1\} = 1,$$

a dla $k \in E$, $k < N$,

$$\begin{aligned} Qh(k) &= \max\left\{ \sum_{j=k+1}^{N} h(j) p_j \prod_{l=k+1}^{j-1} q_l + h(+\infty) \prod_{j=k+1}^{N} q_j,\ \prod_{l=k+1}^{N} q_l \right\} \\ &= \max\left\{ \sum_{j=k+1}^{N} h(j) p_j \prod_{l=k+1}^{j-1} q_l,\ \prod_{l=k+1}^{N} q_l \right\}. \end{aligned}$$

Stąd

$$Q^n r(N) = 1 r(N), \quad Q^n r(+\infty) = 0 = r(+\infty), \qquad n = 0, 1, 2, \ldots,$$

a dla $k < N$,

$$
\begin{aligned}
Qr(k) &= \max\left\{\sum_{j=k+1}^{N} r(j)p_j \prod_{l=k+1}^{j-1} q_l,\ \prod_{l=k+1}^{N} q_l\right\} \\
&= \max\left\{\sum_{j=k+1}^{N} \prod_{l=j+1}^{N} q_l p_j \prod_{l=k+1}^{j-1} q_l,\ \prod_{l=k+1}^{N} q_l\right\} \\
&= \prod_{l=k+1}^{N} q_l \max\left\{\sum_{j=k+1}^{N} \frac{p_j}{q_j}, 1\right\} \\
&= r(k) \max\left\{\sum_{j=k+1}^{N} \frac{p_j}{q_j}, 1\right\}.
\end{aligned}
$$

Niech

$$
t_N := \max\left\{k \le N\colon \sum_{j=k+1}^{N} \frac{p_j}{q_j} \ge 1\right\}.
$$

Dla dowolnych $n > 0$ i $k \in \{1, \ldots, N\}$ mamy więc

$$
Q^n r(k) = r(k) \quad \text{wtedy i tylko wtedy gdy } k > t_N,
$$

Policzmy teraz moment Markowa

$$
\widehat{\tau}_N = \inf\left\{n \le N\colon Q^{N-n} r(X_n) = r(X_n)\right\}.
$$

Mamy

$$
\widehat{\tau}_N = \inf\left\{n \le N\colon X_n \ge t_N\right\}.
$$

Policzmy teraz $Q^N r(1)$. Jeżeli $1 \le t_N$, to

$$
Q^N r(1) = r(1) \left(\sum_{j=k+1}^{N} \frac{p_j}{q_j}\right)^N =
$$

Jeżeli $1 > t_N$, to

$$
Q^N r(1) = r(1) =
$$

8.1 Zastosowania

8.1.1 Problem sekretarki

W klasycznym problemie sekretarki mamy znaną i skończoną liczbę N kandydatek. Osoba rekrutująca egzaminuje kolejno kandydatki. Raz odrzucona kandydatka nie może być zatrudniona. Rekruter podejmuję decyzję na podstawie rekordów kandydatek dotychczas egzaminowanych. Celem rekrutera jest wybór najlepszej kandydatki.

Dokładnie niech ξ_k oznaczą kwalifikacje k-tej kandydatki. Zakładamy, że (ξ_k) jest ciągiem niezależnych zmiennych losowych o rozkładzie jednostajnym na odcinku $[0,1]$. Niech $X_0 = 1$,

$$X_n = \min\left\{k \le N\colon \xi_k > \xi_{X_{n-1}}\right\}, \qquad n = 1, 2, \ldots,$$

przy czym przyjmijmy, że $\min\emptyset = +\infty$. Niech $(\mathfrak{F}_n)$ oznacza filtrację generowaną przez proces (X_n). Niech $\varSigma_N$ oznacza klasę momentów stopu τ względem $(\mathfrak{F}_n)$ takich, że $\tau \le N$. Celem jest maksymalizacja funkcjonału

$$J(\tau) = \mathbb{P}\left\{X_\tau < +\infty, X_{\tau+1} = +\infty\right\}, \qquad \tau \in \varSigma_N.$$

Oznaczmy przez A_k zdarzenie, że k-ta kandydatka jest lepsza od wszystkich poprzednich. Mamy

$$\begin{aligned}\mathbb{P}(A_k) &= \int_0^1 d\xi_k \int_0^{\xi_k} d\xi_1 \ldots \int_0^{\xi_k} d\xi_{k-1}\\ &= \int_0^1 \xi_k^{k-1} d\xi_k = \frac{1}{k}.\end{aligned}$$

Następnie dla $k < j$,

$$\begin{aligned}\mathbb{P}(A_k \cap A_j) &= \int_0^1 d\xi_j \int_0^{\xi_j} d\xi_k \ldots \int_0^{\xi_j} d\xi_{j-1} \int_0^{\xi_k} d\xi_1 \ldots \int_0^{\xi_k} d\xi_{k-1}\\ &= \int_0^1 d\xi_j \int_0^{\xi_j} \xi_k^{k-1} d\xi_k \xi_j^{j-k-1} = \int_0^1 d\xi_j \frac{1}{k}\xi_j^k \xi_j^{j-k-1}\\ &= \frac{1}{k}\frac{1}{j} = \mathbb{P}(A_j)\,\mathbb{P}(A_k).\end{aligned}$$

Stąd założenia twierdzenia Brussa są spełnione. Mamy

$$p_k = \frac{1}{k}, \qquad q_k = 1 - \frac{1}{k} = \frac{k-1}{k}, \qquad r_k = \frac{1}{k-1},$$

przy czym $r_1 = \infty$. Tak więc

$$\begin{aligned}t_N^* &= \max\left\{k = 1, 2, \ldots, N\colon \sum_{j=k}^{N} \frac{1}{j-1} = \sum_{j=k-1}^{N-1} \frac{1}{j} \ge 1\right\}\\ &= \max\left\{k = 1, 2, \ldots, N\colon \sum_{j=k}^{N-1} \frac{1}{j} \le 1\right\},\end{aligned}$$

a optymalnym momentem zatrzymania jest

$$\widehat{\tau}_N = \inf\left\{n \le N\colon X_n \ge t_N^*\right\}.$$

Mamy również

$$J(\widehat{\tau}_N) = \prod_{j=t_N^*}^{N} q_j \sum_{k=t_N^*}^{N} r_k = \prod_{j=t_N^*}^{N} \frac{j-1}{j} \sum_{k=t_N^*}^{N} \frac{1}{k-1} = \frac{t_N^*-1}{N} \sum_{k=t_N^*}^{N} \frac{1}{k-1}.$$

Ponieważ

$$\sum_{k=t_N^*}^{N} \frac{1}{k-1} = \frac{1}{t_N^*-1} + 1 - \frac{1}{N} \to 1,$$

mamy

$$\lim_{N\to\infty} J(\widehat{\tau}_N) = \lim_{N\to\infty} \frac{t_N^*-1}{N}.$$

Ponieważ

$$1 \approx \sum_{k=t_N^*}^{N} \frac{1}{k} \approx \log N - \log t_N^*,$$

otrzymujemy

$$t_N^* \approx \mathrm{e}^{\log N - 1} = N\mathrm{e}^{-1}.$$

W rezultacie

$$\lim_{N\to\infty} J(\widehat{\tau}_n) = \lim_{N\to\infty} \frac{\mathrm{e}^{-1}N-1}{N} = \mathrm{e}^{-1}.$$

8.1.2 Sekretarki przychodzące w grupach

Rozważmy modyfikacje problemu sekretarki polegającą na tym, że kandydatki przychodzą na rozmowę kwalifikacyjną w grupach $m_1, m_2, \ldots, m_N$-kandydatek. Oznaczmy przez A_k zdarzenie, że najlepsza kandydatka w k-tej grupie jest lepsza od wszystkich dotychczasowych kandydatek. Niech $M_k = m_1 + \ldots + m_k$. Mamy

$$\mathbb{P}(A_k) = m_k \int_0^1 d\xi \int_0^{\xi} d\xi_1 \ldots \int_0^{\xi} d\xi_{M_k-1} = \frac{m_k}{M_k} = \frac{m_k}{\sum_{j=1}^{k} m_j}.$$

Łatwo pokazać, że zdarzenia (A_n) są niezależne. Stąd, z twierdzenia Brussa, optymalną regułą stopu jest zatrudnienie najlepszej kandydatki w grupie

$$\widehat{\tau} = \inf\{k \le N \colon k \ge t^* \text{ oraz } A_k \text{ zachodzi}\},$$

gdzie $\inf \emptyset = N$ oraz

$$\tau^* = \max\left\{1, \max\left\{k = 1, 2, \ldots, N \colon \sum_{j=k}^{N} \frac{m_k}{\sum_{j=1}^{k-1} m_j} \ge 1\right\}\right\}.$$

8.1.3 Problem ostatniego skoku

Niech Π będzie procesem Poissona z intensywnością γ. Ustalmy skończony horyzont czasowy T i filtrację $(\mathfrak{F}_t)$ generowaną przez Π. Niech Σ_T oznacza przestrzeń momentów Markowa, które zapadają do chwili T. Szukamy $\tau \in \Sigma_\tau$, które maksymalizuje

$$J_T(\tau) = \mathbb{P}\left(\Pi(t) - \Pi(t-) = 0 \quad \text{dla } t \in (\tau, T]\right\}. \tag{8.1}$$

czyli momentu ostatniego skoku procesu Poissona przed czasem T.

W tym celu rozważmy ciąg podziałów $[0,T]$,

$$[0,T] = \{0\} \cup (t_0^n, t_1^n] \cup \ldots \cup (t^{n_{n-1}}, T],$$

gdzie $t_k^n = kT/n$. Niech $\mathfrak{F}_k^n = \sigma\left(\Pi(s)\colon s \le kT/n\right)$. Niech A_k^n oznacza zdarzenie, że w predziale $(t_{k-1}^n, t_k^n]$ proces Poissona miał co najmniej jeden skok. Oczywiście dla ustalonego n, zdarzenia (A_k^n) są niezależne. Ponadto

$$\begin{aligned}\mathbb{P}\left(A_k^n\right) &= 1 - \mathbb{P}\left(\Pi(t_k^n) - \Pi(t_{k-1}^n) = 0\right) = 1 - \mathbb{P}\left(\Pi(t_k^n - t_{k-1}^n) = 0\right)\\ &= 1 - \mathbb{P}\left(\Pi(T/n) = 0\right) = 1 - e^{-\gamma T/n}.\end{aligned}$$

Niech

$$\begin{aligned}t_n^* &= \max\left\{1, \max\left\{k = 1, 2, \ldots, N\colon (n-k+1)\frac{1 - e^{-\gamma T/n}}{e^{-\gamma T/n}} \ge 1\right\}\right\}\\ &= \max\left\{1, \left[n - 1 - \frac{e^{\gamma T/n}}{1 - e^{-\gamma T/n}}\right]\right\},\end{aligned}$$

gdzie $[a]$ oznacza część całkowitą liczby a.

Z twierdzenia Brussa dla każdego n, moment $\widehat{\tau}^n$, który maksymalizuje prawdopodobieństwo

$$\mathbb{P}\left(\Pi(\tau T/n) - \Pi((\tau-1)T/n) > 0, \quad \Pi(T) - \Pi(\tau T/n) = 0\right)$$

po wszystkich momentach $\tau \in \{1, 2, \ldots, n\}$ względem filtracji $(\mathfrak{F}_k^n)$ dany jest wzorem

$$\widehat{\tau}^n := \min\left\{k \le n\colon \Pi(kT/n) - \Pi((k-1)T/n) > 0, \qquad k \ge t_n^*\right\},$$

gdzie przyjmujemy, że $\min \emptyset := n$.

Stąd (co najmniej kandydatem na optymalny moment zatrzymania dla problemu (8.1) jest

$$\widehat{\tau} = \lim_{n\to\infty} \widehat{\tau}^n \frac{T}{n} = \left(T - \frac{1}{\gamma}\right)^+.$$

Co więcej odpowiadające mu prawdopodobieństwo wynosi

$$\lim_{n\to\infty}\prod_{j=t_n^*}^{n} e^{-\gamma t/n}\sum_{k=t_n^*}^{n}\frac{1-e^{-\gamma T/n}}{e^{-\gamma T/n}} = \lim_{n\to\infty} e^{-\frac{\gamma T(n-t_n^*+1)}{n}}(n-t_n^*+1)\left(e^{\gamma T/n}-1\right)$$
$$= \lim_{n\to\infty} e^{-\gamma T\left(1-\frac{t_n^*}{n}\right)}T\gamma\left(1-\frac{t_n^*}{n}\right)$$
$$= \gamma\left(T-\left(T-\frac{1}{\gamma}\right)^+\right)e^{-\gamma\left(T-\left(T-\frac{1}{\gamma}\right)^+\right)}.$$

W szczególności, jeżeli $T > \frac{1}{\gamma}$, to prawdopodobieństwo wynosi e^{-1}.

8.1.4 Problem wynajmu apartamentu

W problemie wynajmu apartamentu mamy ustalony czas T na wynajem. Oferty przychodzą w momentach skoków procesu Poissona z intenzywnością γ. Ich wysokość modelujemy niezależnymi zmiennymi o rozkładzie jednostajnym na odcinku $[0,1]$. Szukamy momentu stopu (przyjęcia oferty), który maksymalizuje prawdopodobieństwo, że przyjęta oferta jest najkorzystniejsza.

8.2 Problem sterowania ze skończonymi przestrzeniami stanów i sterowań

Załóżmy, że przestrzenie stanów i sterowań są następujące:

$$E = \{x_1,\ldots,x_d\},$$
$$U = \{u_1,\ldots,u_m\}.$$

Dla $x_i \in E$ niech

$$\mathcal{U}(x_i) = \left\{u_1^i,\ldots,u_{m_i}^i\right\}$$

będzie zbiorem parametrów sterujących dopuszczalnych gdy proces jest w stanie x_i.

Niech $[P^u;\ u\in U]$ będzie rodziną prawdopodobieństw przejścia. Możemy P^u identyfikować z macierzą $[P^u_{i,j}] \in M(d\times d)$;

$$P^u_{i,j} = \mathbb{P}\left\{X^\pi_{n+1} = x_j | u_n = u,\ X^\pi_n = x_i\right\}.$$

Oczywiście $P^u_{i,j}\in[0,1]$ oraz $\sum_{j=1}^d P^u_{i,j} = 1$ dla każdego i.

Mamy zadane koszty/zyski bierzące

$$g(x,u), \qquad x\in E,\ u\in U, \quad q(x,u)\geq 0$$

oraz współczynnik dyskonta $\gamma\in(0,1]$. Niech

$$J(X_0,\pi) = \mathbb{E}\sum_{n=0}^{\infty}\gamma^n q(X_n^\pi, u_n(X_0^\pi,\ldots,X_n^\pi)).$$

Ponieważ $E\times U$ jest zbiorem skończonym więc q jest funkcją ograniczoną. Stąd mamy

Lemat 8.1 *Jeżeli $\gamma < 1$, to zboór*

$$\{J(X_0, \pi), \quad \pi \text{ strategia dopuszczalna, } X_0 \text{ zmienna losowa w } E,\}$$

jest ograniczony.

Przypomnijmy, że

$$\mathcal{A}h(x) = \min(\text{odp.}\max)\left\{q(x,u) + \gamma P^u h(x)\colon u \in \mathcal{U}(x)\right\}.$$

Ponieważ E i U są skończone mamy następujący fakt.

Lemat 8.2 *Niech $v_n = \mathcal{A}^n(0)$. Wówczas*

$$0 \le v_\infty := \lim_{n\to\infty} v_n < \infty.$$

Propozycja 8.1 *Przypomnijmy, że jeżeli J jest funkcjonałem zysku to $v_\infty = \mathcal{A}v_\infty$. Ze skończoności zbioru U wynika, że istnieje $u_\infty\colon E \mapsto U$, takie że $u_\infty(x) \in \mathcal{U}(x)$ dla $x \in E$ oraz*

$$v_\infty(x) = q(x, u_\infty(x)) + \gamma P^{u_\infty(x)} v_\infty(x), \qquad x \in E.$$

Ponieważ v_∞ jest ograniczone, więc v_∞ jest funkcją wartości, a

$$\widehat{\pi} = (\widehat{u}_n),$$

gdzie $\widehat{u}_n(x_0, \ldots, x_n) = u_\infty(x_n)$, jest optymalną strategią.

Lemat 8.3 *Rozważmy przypadek funkcjonału zysku. Załóżmy, że v spełnia równanie $v = \mathcal{A}v$. Wówczas $v = v_\infty$.*

Dowód Załóżmy, że v spełnia równania $v = \mathcal{A}v$. Pokażemy, że v jest funkcją wartości dla naszego problemu. Z jedyności funkcji wartości z tego, że v_∞ jest funkcją wartości wyniknie, że $v = v_\infty$. Niech

$$J_N(X_0, \pi) = \mathbb{E}\left\{\sum_{n=0}^{N-1} \gamma^n q(X_n^\pi, u_n(X_0^\pi, \ldots, X_n^\pi)) + \gamma^N v(X_N^\pi)\right\}.$$

Wówczas, ponieważ v jest ograniczone i $\gamma < 1$, zachodzi

$$J_N(X_0, \pi) \to J(X_0, \pi).$$

Niech π_v będzie optymalną strategią dla funkcjonału zysku J_N. Ponieważ $\mathcal{A}v = v$, więc strategia π_v jest stacjonarna i wspolna dla wszystkich horyzontów czasowych N. Stąd dla każdej strategii dopuszczalnej π;

$$J_N(X_0, \pi) \le J_N(X_0, \pi_v) = \mathbb{E}\, v(X_0) \to J(X_0, \pi_v).$$

W rezultacie

$$J(X_0, \pi) \le J(X_0, \pi_V) = \mathbb{E}\, v(X_0). \qquad \square$$

Propozycja 8.2 *Załóżmy, że J jest funkcjonałem kosztu. Wówczas v_∞ jest funkcją wartości oraz jedynym skończonym rozwiązaniem równania $v = \mathcal{A}v$.*

9

Sterowanie ergodyczne

9.1 Wstęp

Zakładamy, że proces sterowany X dany jest rekurencyjnie

$$X_{n+1} = F(X_n, u_n, \xi_{n+1}), \qquad n = 0, 1, \ldots,$$

gdzie X_0 dany element mierzalnej przestrzeni stanów $(E, \mathcal{E})$, $\{\xi_n\}$ jest ciągiem iid o wartościach w mierzalnej przestrzeni $(S, \mathcal{S})$, $(U, \mathcal{U})$ jest mierzalną przestrzenią sterowań, a F jest mierzalnym odwzorowaniem $E \times U \times S$ w E. Jak w rozdziale poświęconym sterowaniu na skończonym przedziale czasowym zakładamy, że dla $x \in E$ określony jest zbiór $U(x) \in \mathcal{U}$ sterowań dopuszczalnych ze stanu x.

Oznaczamy przez X^π proces odpowiadający strategii

$$\pi = (u_0, u_1, \ldots),$$

patrz rozdział poświęcony sterowaniu na skończonym przedziale.

Funkcjonał kosztu jest postaci

$$J_\infty(\pi, X_0) = \limsup_{N \to +\infty} \frac{1}{N} \mathbb{E} \sum_{n=0}^{N-1} q(X_n, u_n),$$

gdzie $q\colon E \times U \to [0, +\infty)$ jest daną funkcją nieujemną mierzalną. Problem optymalizacji tego typu funkcjonałem kosztu nazywamy *sterowaniem z ergodycznym funkcjonałem kosztu*. Uzasadnienie nazwy ergodyczny jest następujące. Często optymalną strategią $\widehat{\pi}$ jest strategia stacjonarna. Strategie π nazywamy *strategią stacjonarną* gdy

$$\pi = (u_0, u_1, \ldots)\,,$$

gdzie

$$u_n(x_0, \ldots, x_n) = u(x_n), \qquad n = 0, 1, \ldots,\; x_0, \ldots, x_n \in E,$$

dla pewnej ustalonej funkcji u. Piszemy wówczas

$$\pi = (u, u, \ldots) = (u(x_0), u(x_1), \ldots).$$

Załóżmy więc, że

$$\widehat{\pi} = (\widehat{u}, \widehat{u}, \ldots)$$

jest stacjonarną strategią optymalną. Niech $\widehat{X} = (\widehat{X}_n)$ będzie procesem odpowiadającym $\widehat{\pi}$ oraz niech $\widehat{P}$ będzie jego operatorem przejścia. Załóżmy, że funkcja q jest ograniczana. Niech

$$\widehat{q}(x) = q\left(x, \widehat{u}(x)\right), \qquad x \in E.$$

Wówczas mamy

$$\begin{aligned} J_\infty(\widehat{\pi}, X_0) &= \limsup_{N\to+\infty} \mathbb{E}\, \frac{\sum_{n=0}^{N-1} q\left(\widehat{X}_n, \widehat{u}(\widehat{X}_n)\right)}{N} \\ &= \limsup_{N\to+\infty} \mathbb{E}\, \frac{\sum_{n=0}^{N-1} \widehat{q}\left(\widehat{X}_n\right)}{N} \\ &= \limsup_{N\to+\infty} \mathbb{E}\, \frac{\sum_{n=0}^{N-1} \widehat{P}^n \widehat{q}(X_0)}{N}. \end{aligned}$$

Stąd, o ile proces $\widehat{X}$ jest ergodyczny, mamy

$$J_\infty(\widehat{\pi}, X_0) = \int_E \widehat{q}(x)\widehat{\mu}(dx),$$

gdzie $\widehat{\mu}$ jest miarą niezmienniczą dla $(\widehat{X}_n)$.

9.2 Równania Bellmana–Howarda

Przypomnijmy, że operatory przejścia dane są wzorami

$$P^u\psi(x) = \mathbb{E}\,\psi\left(F(x, u, \xi_0\right), \qquad u \in U,\ x \in E,\ \psi \in B_b(E).$$

Równania Bellmana–Howarda dla problemu optymalnego sterowania z ergodycznym funkcjonałem kosztu mają następującą postać

$$\begin{aligned} h(x) &= \inf_{u\in U(x)} P^u h(x), \qquad x \in E, \\ h(x) + V(x) &= \mathcal{A}\, V(x), \qquad x \in E, \end{aligned}$$

gdzie

$$\mathcal{A}\, V(x) := \inf_{u\in U(x)} \left(q(x, u) + P^u V(x)\right).$$

Ogólnym twierdzeniem o istnieniu sterowania optymalnego dla problemu z ergodycznym funkcjonałem kosztu jest następujący rezultat.

Twierdzenie 9.1 *(i) Załóżmy, że (h, V) są nieujemnumi mierzalnymi rozwiązaniami równań Bellmana–Howarda. Wówczas dla dowolnej strategii π spełniającej*

$$\lim_{N\to+\infty} \frac{1}{N}\,\mathbb{E}\,V(X_N) = 0 \tag{9.1}$$

mamy

$$J_\infty(\pi, X_0) \geq \mathbb{E}\,h(X_0). \tag{9.2}$$

(ii) Załóżmy ponadto, że istnieje mierzalne odwzorowanie $\widehat{u}\colon E \to U$, takie że $\widehat{u}(x) \in U(x)$ dla $x \in E$ i

$$h(x) = P^{\widehat{v}(x)}h(x), \quad h(x) + V(x) = q(x, \widehat{v}(x)) + P^{\widehat{v}(x)}V(x), \qquad x \in E.$$

Wówczas jeżeli $\mathbb{E}\,V(X_0) < +\infty$ i $\mathbb{E}\,h(X_0) < +\infty$, to dla strategii

$$\widehat{\pi} = (\widehat{u}_0, \widehat{u}_1, \ldots), \qquad \widehat{u}(x_0, \ldots, x_n) = \widehat{v}(x_n),$$

mamy

$$J_\infty(\widehat{\pi}, X_0) = \mathbb{E}\,h(X_0).$$

Tak więc $\widehat{\pi}$ jest strategią optymalną w klasie strategii spełniających (9.1).

Dowód Pokażemy przez indukcje, że dla dowolnej strategii π i dla dowolnego N zachodzi

$$\mathbb{E}\,h(X_N) \geq \mathbb{E}\,h(X_0), \tag{9.3}$$

$$\mathbb{E}\left(\sum_{n=0}^{n-1} q(X_n, u_n) + V(X_n)\right) \geq \mathbb{E}\left(V(X_0) + N\,h(X_0)\right). \tag{9.4}$$

Dla $N = 0$ mamy (9.3) i (9.4). Teraz zakładając (9.3) i (9.4) dla $0, 1, \ldots, N$ otrzymujemy

$$\begin{aligned}
\mathbb{E}\,h(X_{N+1}) &= \mathbb{E}\,\mathbb{E}\left(h(X_{N+1}|X_N\right) \\
&= \mathbb{E}\,P^{u_N(X_0,\ldots,X_n)}h(X_N) \\
&\geq \mathbb{E}\,h(X_N) \geq \mathbb{E}\,h(X_0),
\end{aligned}$$

co dowodzi (9.3). Dla dowodu (9.4) zauważmy, że

$$\begin{aligned}
&\mathbb{E}\left(\sum_{n=0}^{N} q(X_n, u_n) + V(X_{N+1})\right) \\
&= \mathbb{E}\left(\sum_{n=0}^{N-1} g(X_n, u_n) + V(X_N)\right) + \mathbb{E}\left(q(X_{N+1}, u_{N+1})\right) \\
&\qquad + \mathbb{E}\left(V(X_{N+1}) - V(X_N)\right) \\
&\geq \mathbb{E}\left(V(X_0) + N\,h(X_0)\right) + \mathbb{E}\Big(q(X_N, u_N) + P^{U_N(X_0,\ldots,X_N)}V(X_N)\Big) \\
&\qquad - \mathbb{E}\,V(X_N) \\
&\geq \mathbb{E}\left(V(X_0) + Nh(X_0)\right) + \mathbb{E}\,h(X_N) \\
&\geq \mathbb{E}\left(V(X_0) + (N+1)h(X_0)\right).
\end{aligned}$$

Mając (9.4) otrzymujemy

$$\frac{1}{N}\,\mathbb{E}\sum_{n+0}^{N-1} q(X_n, u_n) + \frac{1}{N}\,\mathbb{E}\,V(X_N) \geq \mathbb{E}\left(\frac{1}{N}\,V(X_0) + h(X_0)\right).$$

przechodząc do granicy i korzystając z (9.1) otrzymujemy żądane oszacowanie. Dowód drugiej części wykorzystuje fakt, że dla strategii $\widehat{\pi}$ wszędzie mamy równości. □

Często rozwiązanie równań Bellmana–Howarda jest postaci (λ, V), gdzie λ jest stałą. Wówczas optymalny koszt nie zależy od warunku początkowego X_0.

9.3 Problem liniowo-kwadratowy

Rozważamy przypadek liniowo-kwadratowy, w którym sterowanie i stan zależą od zaburzenia, patrz Rozdział 6.2. To znaczy

$$X_{n+1} = \Psi X_n + \Phi U_n + A(X_n, \xi_n^0) + B(u_n, \xi_n^1) + C\xi_n,$$

gdzie (ξ_n^0), (ξ_n^1), (ξ_n) niezależne zmienne losowe o wartościach w $\mathbb{R}^{r_1}$, $\mathbb{R}^{r_2}$ i $\mathbb{R}^r$, o zerowych średnich i jednostkowych macierzach kowariancji. Ψ, Φ i C są macierzami odpowiednich wymiarów $M(d \times d)$ i $M(d \times m)$ i $M(d \times r)$, a

$$A\colon \mathbb{R}^d \times \mathbb{R}^{r_1} \to \mathbb{R}^d, \qquad B\colon \mathbb{R}^m \times \mathbb{R}^{r_2} \to \mathbb{R}^d$$

są odwzorowaniami dwuliniowymi.

Funkcjonał kosztu jest postaci

$$J_\infty(\pi, X_0) = \limsup_{N\to+\infty} \frac{1}{N}\,\mathbb{E}\sum_{n=0}^{N-1} \left(\langle QX_n, X_n\rangle + \langle Ru_n, u_n\rangle\right),$$

gdzie Q i R są nieujemnie określonymi symetrycznymi macierzami. Dodatkowo zakładamy, że R jest odwracalna.

Szukamy rozwiązań równań Bellmana–Howarda w postaci (λ, V), gdzie

$$V(x) = \langle Kx, x\rangle, \qquad x \in \mathbb{R}^d.$$

z odpowiednio dobraną macierzą symetryczną dodatnio określoną K. Mamy, patrz Rozdział 4.2,

$$\inf_{u\in U}\left(q(x,u) + P^u V(x)\right) = \langle \mathcal{A}_1(K)x, x\rangle + \operatorname{Trace} C^*KC.$$

Ponadto infimum jest osiągalne w punkcie

$$\widehat{u}(x) = \mathcal{B}_1(K)x,$$

gdzie

$$
\begin{aligned}
\mathcal{A}_1(K) &= Q + \mathcal{G}_1(K) + \Psi^* K \left(I + \Phi(R + \mathcal{G}_2(K))^{-1} \Phi^* K\right) \Psi,\\
\mathcal{B}_1(K) &= -\left(R + \Phi^* K \Phi + \mathcal{G}_2(K)\right)^{-1} \Psi^* K \Psi,\\
\langle \mathcal{G}_1(L)x, y\rangle &= \mathbb{E}\, \langle LA(x, \xi_0^1), A(y, \xi_0^1)\rangle\\
\langle \mathcal{G}_2(L)x, y\rangle &= \mathbb{E}\, \langle LB(x, \xi_0^2), B(y, \xi_0^2)\rangle.
\end{aligned}
$$

Tak więc

$$K = \mathcal{A}_1(K), \qquad \lambda = \mathrm{Trace}\, C^* K C.$$

Z interpretacji $\mathcal{A}_1$ wynika następujący fakt.

Lemat 9.1 *Operator $\mathcal{A}_1$ jest monotoniczny w tym sensie, że*

$$\mathcal{A}_1 K \le \mathcal{A}_1 \tilde{K}, \qquad \forall\, K, \tilde{K} \in M_s^+(d) \colon \tilde{K} - K \in M_s^+(d).$$

Mamy następujące twierdzenie.

Twierdzenie 9.2 *Jeżeli istnieje skończone nieujemne rozwiązanie równania $K = \mathcal{A}_1(K)$, to istnieje rozwiązanie $\widehat{K}$ minimalne w tym sensie, że $K - \widehat{K} \ge 0$ dla dowolnej macierzy K spełniającej $\mathcal{A}_1(K) = K$. Niech*

$$\widehat{\lambda} = \mathrm{Trace}\, C^* \widehat{K} C.$$

Wówczas dla dowolnej strategii π,

$$J_\infty(\pi, X_0) \ge \widehat{\lambda}.$$

Ponadto, o ile $\mathbb{E}\, \langle \widehat{K} X_0, X_0\rangle < +\infty$, to dla strategii stacjonarnej

$$\widehat{\pi} = (\widehat{u}, \widehat{u}, \ldots), \qquad \widehat{u}(x) = \mathcal{B}(\widehat{K})(x), \qquad x \in \mathbb{R}^d,$$

zachodzi

$$J_\infty(\widehat{\pi}, X_0) = \widehat{\lambda}.$$

Dowód Dowód powinien wynikać z Twierdzenia 9.1. Ale w tym przypadku trzeba byłoby pokazać, że

$$\lim_{N \to +\infty} \frac{1}{N}\, \mathbb{E}\, V(X_N) = 0, \qquad \mathbb{E}\, V(X_0) < +\infty.$$

Udowodnimy twierdzenie bezpośrednio. W tym celu zauważmy, że

$$
\begin{aligned}
&\frac{1}{N}\, \mathbb{E} \sum_{n=0}^{N-1} \left(\langle Q X_n, X_n\rangle + \langle R u_n, u_n\rangle\right)\\
&\ge \frac{1}{N}\, \mathbb{E} \left(\langle K_N X_0, X_0\rangle + \sum_{n=0}^{N-1} \mathrm{Trace}\, C^* K_n C\right),
\end{aligned}
$$

gdzie $K_n = \mathcal{A}_1^n(0)$. Z monotoniczności $\mathcal{A}_1$, $K_n \uparrow \widehat{K}$, gdzie $\widehat{K}$ jest skończonym rozwiązaniem równania $\mathcal{A}_1(K) = K$. Oczywiście, z monotoniczności $\mathcal{A}_1$ wynika, że $\widehat{K}$ jest rozwiązaniem minimalnym. Stąd

$$\frac{1}{N}\sum_{n=0}^{N-1} \operatorname{Trace} C^* K_n C \to \operatorname{Trace} C^* \widehat{K} C = \widehat{\lambda}.$$

Po przejściu do granicy otrzymujemy więc żądane oszacowanie. Pokazujemy teraz optymalność strategii $\widehat{\pi}$. Mamy

$$\begin{aligned}
&\frac{1}{N}\,\mathbb{E}\left\{\sum_{n=0}^{N-1}\Big(\langle Q\widehat{X}_n, \widehat{X}_n\rangle + \langle R\widehat{U}_n, \widehat{u}_n\rangle\Big) + \langle \widehat{K}\widehat{X}_N, \widehat{X}_N\rangle\right\}\\
&= \frac{1}{N}\,\mathbb{E}\,\langle \widehat{K}X_0, X_x\rangle + \frac{1}{N}\sum_{n=0}^{N-1} \operatorname{Trace} C^* \mathcal{A}_1^n(\widehat{K}) C\\
&= \frac{1}{N}\,\mathbb{E}\,\langle \widehat{K}X_0, X_0\rangle + \widehat{\lambda}.
\end{aligned}$$

Stąd przechodząc do granicy otrzymujemy, żądaną równość

$$J_\infty(\widehat{\pi}, X_0) = \widehat{\lambda}.$$

□

9.4 Skończona przestrzeń stanów

Załóżmy, że przestrzenie stanów E i sterowań U są skończone. Będziemy korzystali z następującego twierdzenia ergodycznego.

Twierdzenie 9.3 *Załóżmy, że $X = (X_n)$ jest jednorodnym łańcuchem Markowa na skończonej przestrzeni stanów*

$$E = \{1, \dots M\}.$$

Niech P będzie macierzą przejścia X i niech

$$L_N := \frac{1}{N}\left(I + P + \ldots + P^{N-1}\right), \qquad N = 1, 2, \ldots.$$

Wówczas istnieje granica

$$\lim_{N\to+\infty} L_N =: L$$

oraz

$$LP = PL = L = L^2.$$

Dowód Niech

$$L_N = \left(l_{i,j}^{(N)}\right), \qquad N = 1, \dots,\ i, j = 1, \dots, M.$$

Wówczas

$$PL_N = L_N P = L_N + \frac{1}{N}\left(P^N - I\right). \tag{9.5}$$

Ponieważ

$$l_{i,j}^{(N)} \geq 0, \qquad \sum_{k=1}^{M} l_{i,j}^{(N)} = 1,$$

mamy

$$L_{i,j}^{(N)} \leq 1, \qquad N = 1, \dots,\ i, j = 1, \dots, M.$$

Stąd istnieje podciąg zbieżny (L_{N_n}). Niech

$$L = \lim_{n \to +\infty} L_{N_n}.$$

Wówczas , z (9.5),

$$PL = LP = L. \tag{9.6}$$

Niech $\tilde{L}$ będzie granicą innego zbieżnego podciągu $(L_{\tilde{N}_n})$. Z (9.6),

$$L_{\tilde{N}_n} L = L L_{\tilde{N}_n} = L.$$

Stąd

$$\tilde{L} L = L \tilde{L} = L.$$

W ten sam sposób pokazujemy, że

$$L\tilde{L} = \tilde{L} L = \tilde{L}.$$

Czyli mamy

$$\tilde{L} = L = L^2.$$

□

Z następującego twierdzenia wynika istnienie optymalnego sterowania stacjonarnego dla przypadku skończonych E i U.

Twierdzenie 9.4 *Dla dowolnej funkcji $q\colon E \times U \to [0, +\infty)$ istnieje optymalna stacjonarna strategia*

$$\widehat{\pi} = (\widehat{u}, \widehat{u}, \dots,).$$

W dowodzie twierdzenia będziemy korzystali z następującego lematu tauberowskiego.

Lemat 9.2 *Dla dowolnego ciągu liczb nieujemnych* (a_n) *zachodzi*

$$\liminf_{N\to+\infty} \frac{a_0+\ldots a_{N-1}}{N} \le \liminf_{\lambda\uparrow 1}(1-\lambda)\sum_{n=0}^{+\infty} a_n\lambda^n$$
$$\le \limsup_{\lambda\uparrow 1}(1-\lambda)\sum_{n=0}^{+\infty} a_n\lambda^n \le \limsup_{N\to+\infty} \frac{a_0+\ldots a_{N-1}}{N}.$$

Dowód Niech

$$S_n := \sum_{j=0}^{n} a_j, \qquad n=0,1,\ldots,$$
$$a(\lambda) := \sum_{j=0}^{+\infty} a_n\lambda^n, \qquad \lambda\in[0,1).$$

Przyjmując $S_{-1}=0$ otrzymujemy

$$a(\lambda) = \sum_{n=0}^{+\infty}(S_n - S_{n-1})\lambda^n = (1-\lambda)\sum_{n=0}^{+\infty} S_n\lambda^n.$$

Stosując tą równość do ciągu stałego $a_n=1$ otrzymujemy

$$(1-\lambda)^{-2} = (1-\lambda)^{-1}\sum_{n=0}^{+\infty}\lambda^n = \sum_{n=0}^{+\infty}(n+1)\lambda^n.$$

Teraz, jeżeli dla liczb r i n_0 zachodzi

$$\frac{S_n}{n+1} \le r, \qquad n\ge n_0,$$

to

$$\begin{aligned}(1-\lambda)a(\lambda) &= (1-\lambda)^2\sum_{n<n_0} S_n\lambda^n + (1-\lambda)^2\sum_{n\ge n_0} S_n\lambda^n\\ &\le (1-\lambda)^2\sum_{n<n_0} S_n\lambda^n + r(1-\lambda)^2\sum_{n\ge n_0}(n+1)\lambda^n\\ &\le (1-\lambda)^2\sum_{n<n_0} S_n\lambda^n + r.\end{aligned}$$

Stąd

$$\limsup_{\lambda\uparrow 1}(1-\lambda)a(\lambda) \le r,$$

a więc

$$\limsup_{\lambda\uparrow 1}(1-\lambda)a(\lambda) \le \limsup_{n\to+\infty}\frac{S_n}{n+1}.$$

Podobnie, jeżeli
$$\frac{S_n}{n+1} \geq r \qquad \text{dla } n \geq n_0,$$
to
$$\begin{aligned}(1-\lambda)^2 \sum_{n\geq n_0} S_n\lambda^n &= (1-\lambda)^2 \sum_{n\geq n_0} \frac{S_n}{n+1} n+1\lambda^n \\ &\geq r(1-\lambda)^2 \sum_{n\geq n_0} (n+1)\lambda^n \\ &\geq r\left(1-(1-\lambda)^2 \sum_{n<n_0} (n+1)\lambda^n\right).\end{aligned}$$
Stąd
$$\liminf_{\lambda\uparrow 1}(1+\lambda)a(\lambda) \geq \liminf_{n\to+\infty} \frac{S_n}{n+1}. \qquad \square$$

Dowód Twierdzenia 9.4 Dla $\gamma \in (0,1)$ niech
$$\mathcal{A}_\gamma\psi(x) = \inf_{u\in U(x)} \left(q(x,u) + \gamma P^u\psi(x)\right), \qquad x \in E,\ \psi \in B_b(E).$$
Pokażemy, że odwzorowania
$$\mathcal{A}_\gamma \colon B_b(E) \to B_b(E), \qquad \gamma \in (0,1),$$
są (ścisłymi) kontrakcjami. Dokładnie mamy
$$||\mathcal{A}_\gamma\psi - \mathcal{A}_\gamma\phi||_{B_b(E)} := \max_{x\in E} |\mathcal{A}_\gamma\psi(x) - \mathcal{A}_\gamma\phi(x)| \leq \gamma||\psi-\phi||_{B_b(E)}.$$
Istotnie,
$$\begin{aligned}&|\mathcal{A}_\gamma\psi(x) - \mathcal{A}_\gamma\phi(x)| \\ &= \left|\min_{u\in U(x)} \left(q(x,u) + \gamma P^u\psi(x)\right) - \min_{u\in U(x)} \left(q(x,u) + \gamma P^u\phi(x)\right)\right| \\ &\leq \max_{u\in U(x)} |\gamma P^u\psi - \gamma P^u\phi(x)| \\ &\leq \gamma\, ||\psi-\phi||_{B_b(E)}.\end{aligned}$$
Z twierdzenia Banacha o punkcie stałym wynika, że dla dla każdego $\gamma \in (0,1)$ istnieje dokładnie jedna funkcja $V_\gamma \in B_b(E)$ spełniająca
$$\mathcal{A}_\gamma V_\gamma = V_\gamma.$$
Ponieważ $\mathcal{A}_\gamma$ przeprowadza funkcje nieujemne w nieujemne, mamy $V_\gamma \geq 0$, $\gamma \in (0,1)$. Ponieważ, E i U są skończone, dla każdego $\gamma \in (0,1)$ istnieje mierzalne $\widehat{u}_\gamma \colon E \to U$, takie że

$$\widehat{u}_\gamma(x) \in U(x), \qquad \mathcal{A}_\gamma V_\gamma(x) = q(x, \widehat{u}_\gamma(x)) + \gamma P^{\widehat{u}_\gamma(x)} V_\gamma(x), \qquad x \in E.$$

Z Twierdzenia 3.1 strategia stacjonarna

$$\widehat{\pi}_\gamma = (\widehat{u}_\gamma, \widehat{u}_\gamma, \ldots)$$

optymalizuje funkcjonał kosztu

$$J^\gamma(\pi, X_0) = \mathbb{E} \sum_{n=0}^{+\infty} \gamma^n q(X_n, u_n).$$

Ponieważ istnieje tylko skończona liczba różnych funkcji z E w U, istnieje ciąg $\gamma_j \uparrow 1$, taki że

$$\widehat{u}_{\gamma_j} = \widehat{u}_{\gamma_i}, \qquad \forall\, j, i.$$

Pokażemy, że strategia stacjonarna

$$\widehat{\pi} = (\widehat{u}, \widehat{u}, \ldots), \qquad \text{z} \qquad \widehat{u} := \widehat{u}_{\gamma_1} = \widehat{u}_{\gamma_2} = \ldots,$$

jest optymalna dla problemu ergodycznego. Istotnie, Niech π będzie dowolną strategią. Oznaczmy przez $\widehat{X}$ proces odpowiadający $\widehat{\pi}$. Wówczas dla dowolnego $\gamma = \gamma_j$ mamy

$$(1-\gamma)\, \mathbb{E} \sum_{n=0}^{+\infty} \gamma^n q(X_n, u_n) \geq (1-\gamma)\, \mathbb{E} \sum_{n=0}^{+\infty} \gamma^n q\left(\widehat{X}_n, \widehat{u}(\widehat{X}_n)\right).$$

Z lematu tauberowskiego (Lemat 9.2) mamy

$$\limsup_{N\to+\infty} \frac{1}{N}\, \mathbb{E} \sum_{n=0}^{N-1} q(X_n, u_n) \geq \limsup_{N\to+\infty} \frac{1}{N}\, \mathbb{E} \sum_{n=0}^{N-1} q\left(\widehat{X}_n, \widehat{u}_n(\widehat{X}_n)\right).$$

Z twierdzenia ergodycznego (Twierdzenie 9.3) istnieje granica

$$\begin{aligned}
&\lim_{N\to+\infty} \frac{1}{N}\, \mathbb{E} \sum_{n=0}^{N-1} q\left(\widehat{X}_n, \widehat{u}_n(\widehat{X}_n)\right) \\
&= \lim_{N\to+\infty} \frac{1}{N}\, \mathbb{E} \sum_{n=0}^{N-1} \widehat{P}^n \widehat{q}(X_0) \\
&= \mathbb{E}\, \widehat{L}\widehat{q}(X_0),
\end{aligned}$$

gdzie $\widehat{P}$ jest operatorem przejścia $\widehat{X}$,

$$\widehat{q}(x) = q\left(x, \widehat{u}(x)\right), \qquad x \in E,$$

a $\widehat{L}$ jest operatorem granicznym występującym w twierdzeniu ergodycznym. Stąd otrzymujemy żądany wniosek

$$J_\infty(\pi, X_0) \geq J_\infty(\widehat{\pi}, X_0). \qquad \square$$

9.5 Algorytm Howarda

9.5.1 Rezultaty wstępne

Niech X będzie jednorodnym procesem Markowa na mierzalnej przestrzeni stanów $(E, \mathcal{E})$. Oznaczmy przez P operator przejścia dla X. Dla dowolnej miary μ na $(E, \mathcal{E})$ definiujemy miare $P^*\mu$ na $(E, \mathcal{E})$ wzorem

$$\langle P * \mu, \psi \rangle := \int_E \psi(x) P^*\mu(dx) = \langle \mu, P\psi \rangle, \qquad \forall\, \psi \in B_b(E).$$

Przypomnijmy, że *miara niezmiennicza dla procesu Markowa* na E z operatorem przejścia P to dowolna miara probabilistyczna μ na $(E, \mathcal{E})$ spełniająca $P^*\mu = \mu$. Oczywiście gdy $E = \{1, \ldots, d\}$ jest skończone, to $P \in M(d \times d)$ a każda miara niezmiennicza może być identyfikowana z wektorem $\mu \in \mathbb{R}^d$, takim że

$$P^*\mu = \mu, \qquad \mu = [\mu_1, \ldots, \mu_d],\ \mu_i \geq 1,\ i = 1, \ldots, d,\ \sum_{i=1}^{d} \mu_i = 1.$$

Będziemy potrzebowali następującego twierdzenia.

Twierdzenie 9.5 *Załóżmy, że P jest macierzą przejścia dla jednorodnego łańcucha Markowa na skończonej przestrzeni stanów*

$$E = \{1, \ldots, d\}.$$

Niech

$$L = \lim_{N \to +\infty} \frac{1}{N} \left(I + P + \ldots + P^{N-1} \right).$$

(i) Wektor $\mu \in \mathbb{R}^d$ spełnia $P^*\mu = \mu$ wtedy i tylko wtedy gdy spełnia $L^*\mu = \mu$. Ponadto zbiór wszystkich rozwiązań równania $P^*\mu = \mu$ jest równy $\operatorname{Im} L^*$.
(ii) Wektor $h \in \mathbb{R}^d$ spełnia $Ph = h$ wtedy i tylko wtedy gdy spełnia $Lh = h$. Ponadto zbiór wszystkich rozwiązań równania $Ph = h$ jest równy $\operatorname{Im} L$.
(iii) Dowolna miara niezmiennicza μ dla X jest postaci

$$\mu = \mu^1 \gamma_1 + \ldots + \mu^d \gamma_d$$

gdzie μ^i jest i-tą kolumną macierzy L^*, a

$$\gamma_i \geq 0,\ i = 1, \ldots, d, \qquad \sum_{i=1}^{d} \gamma_i = 1.$$

(iv) Istnieje dokładnie jedna miara niezmiennicza dla X wtedy i tylko wtedy gdy wszystkie rozwiązania równania $Ph = h$ są wektorami postaci λI, $\lambda \in \mathbb{R}$, I macierz jednostkowa w $\mathbb{R}^d$.

Dowód (i) Jeżeli $P^*\mu = \mu$ to $L_N^*\mu = \mu$, gdzie

$$L_N = \frac{1}{N}\left(I + P + \ldots + P^{N-1}\right)$$

Stąd po przejściu do granicy otrzymujemy $L^*\mu = \mu$. Niech μ spełnia $L^*\mu = \mu$. Wówczas, z twierdzenia ergodycznego

$$P^*\mu = P^*L^*\mu = (LP)^*\mu = (PL)^*\mu = L^*\mu = \mu.$$

Oczywiście jeżeli $P^*\mu = \mu$ to z udowodnionej równoważności $L^*\mu = \mu$, a więc $\mu \in \operatorname{Im} L^*$. Niech $\mu = L^*\xi$ dla $\xi \in \mathbb{R}$. Wówczas

$$L^*\mu = L^*L^*\xi = (LL)^*\xi = L^*\xi = \mu.$$

co kończy dowód (i) Dowód (ii) jest analogiczny i pozostawiamy go czytelnikowi.

Przechodzimy do dowodu (iii). Jeżeli μ ma postać opisaną w (iii) to $\mu = L^*\gamma$ dla wektora γ o współrzędnych (γ_i). Stąd

$$P^*\mu = P^*L^*\gamma = (LP)^*\gamma = L^*\gamma = \mu.$$

Odwrotnie, jeśli $P^*\mu = \mu$, to z piewszej części $\mu = L^*\mu$ i dowód (iii) jest zakończony.

Z (iii) wynika, że istnieje dokładnie jedna miara niezmiennicza wtedy i tylko wtedy gdy L^* ma jednakowe kolumny. Stąd rozwiązania równania $Lh = h$ są postaci λI i (iv) wynika z (ii). $\square$

Twierdzenie 9.6 *Niech X będzie jednorodnym procesem Markowa X na skończonej przestrzeni stanów $E = \{1, \ldots, d\}$. Niech P będzie macierzą przejścia dla X. Niech $q \in \mathbb{R}^d$. Rozważmy następujące* liniowe równanie Bellmana–Howarda

$$h = Ph, \qquad h + V = q + PV. \tag{9.7}$$

Jeżeli dla X istnieje tylko jedna miara niezmiennicza μ to istnieje para $(h, V) \in \mathbb{R}^d \times \mathbb{R}^d$ spełniająca (9.7). *Ponadto*

$$h = \langle \mu, q \rangle I,$$

czyli jest wyznaczona jednoznacznie oraz dla dowolnych rozwiązań

$$(h, V_1) \qquad \text{i} \qquad (h, V_2)$$

zachodzi

$$V_1 - V_2 = \lambda I,$$

dla jakiejś stałej λ.

Dowód Z Twierdzenie 9.5(iv), wszystkie rozwiązania równania $Ph = h$ są postaci γI. Tak więc (9.7) sprowadza się do postaci

$$\gamma I + V = q + PV.$$

Dla dowolnego N zachodzi więc

$$\frac{1}{N}\sum_{n=0}^{N-1}(\gamma I + P^n V) = \frac{1}{N}\sum_{n=0}^{N-1}P^n q + \frac{1}{N}\sum_{n=0}^{N-1}P^{n+1}V.$$

Czyli, odejmując stronami otrzymujemy

$$\gamma I + \frac{V}{N} = L_N q + \frac{P^N V}{N}.$$

Przechodząc do granicy otrzymujemy

$$\gamma I = Lq.$$

Macierz L ma jednakowe wiersze, bo L^* ma jednakowe kolumny. Ponadto wiersze L są współrzędnymi miary niezmienniczej μ. Stąd

$$Lq = \langle \mu, q\rangle I.$$

Pokazaliśmy więc, że pierwszy wektor h dowolnego rozwiązania (h, V) równania (9.7) musi być postaci

$$h = \langle \mu, q\rangle I.$$

Pozostaje do wykazania istnienie wektora V i jego jedyność modulo wektor cI, $c \in \mathbb{R}$. Do wykazania istnienia należy pokazać, że istnieje V spełniające

$$(I - P)V = q - \langle \mu, q\rangle I = (I - L)q.$$

Czyli zawieranie się obrazów

$$\mathrm{Im}\,(I - L) \subseteq \mathrm{Im}\,(I - P).$$

Wynika to z następującego rozumowania. Jeżeli inkluzja niezachodziła by to istniałby wektor $y \in \mathbb{R}$ taki, że

$$\langle (I - P)x, y\rangle = 0, \qquad \forall\, x \in \mathbb{R}^d$$

oraz istniałby $\widehat{x} \in \mathbb{R}^d$, taki że

$$\langle (I - L)\widehat{x}, y\rangle \neq 0.$$

Z pierwszego warunku mamy $P * y = y$. Z Twierdzenia 9.5(i), $L^* y = y$. Czyli

$$\langle (I - L)\widehat{x}, y\rangle = \langle \widehat{x}, (I - L^*)y\rangle = 0$$

co prowadzi do sprzeczności. Istnieje więc rozwiązanie (h, V) równania (9.7) i pierwszy wektor ma postać $\langle \mu, q\rangle I$. Niech (h, V_1) i (H, V_2) będą dwoma rozwiązaniami. Wówczas, z dowodu poprzednich elementów twierdzenia

$$V_1 - V_2 = P(V_1 - V_2) = \lambda I$$

dla pewnej $\lambda \in \mathbb{R}$. $\square$

9.5.2 Algorytm

Algorytm Howarda służy do znajdywania optymalnej strategii stacjonarnej dla problemu z kosztem ergodycznym. Polega na sukcesywnym polepszaniu strategii.

Oznaczmy przez V zbiór wszystkich dopuszczalnych sterowań, to znaczy

$$V = \{u\colon E \to U\colon u(x) \in U(x),\ \forall\, x \in E\}\,.$$

Pisząc, dla $u \in V$,

$$(u) := (u, u, \ldots)$$

będziemy utożsamiali strategie stacjonarną ze sterowaniem u. Odpowiadający strategii (u) łańcuch Markowa oznaczmy X^u. Będziemy zakładali, że dla każdej $u \in V$, X^u ma dokładnie jedną miarę niezmienniczą μ_u. Przez P^u będziemy oznaczali macierz przejścia dla X^u.

Oczywiście dla $u \in V$ mamy

$$J_\infty((u), X_0) = \lambda_u := \langle \mu_u, q(\cdot, (u(\cdot))\rangle.$$

Definiujemy odwzorowanie $\tau\colon V \to V$ w następujący sposób. Rozwiązujemy liniowe równanie Bellmana–Howarda, patrz Twierdzenie 9.6,

$$\lambda_u + V_u(x) = q(x, u(x)) + P^{u(x)} V_u(x), \qquad x \in E.$$

Ponieważ V_u jest wyznaczone jednoznacznie z dokładnością do stałej V_u będzie wyznaczona jednoznacznie gdy założymy, że $V_u(1) = 0$. Sprawdzamy, teraz czy (λ_u, V) spełnia równanie (nieliniowe) Bellmana–Howarda, to znaczy czy

$$\lambda_u + V_u(x) = \mathcal{P} V_u(x), \qquad x \in E, \tag{9.8}$$

gdzie

$$\mathcal{P} V_u(x) := \min_{v \in U(x)} \{q(x, v) + P^v V_u(x)\}\,, \qquad x \in E.$$

Zauważmy, że ponieważ $P^v 1 = 1$, $v \in U$, to zawsze mamy

$$\lambda_u = \min_{v \in U(x)} P^v \lambda_u(x), \qquad x \in E.$$

Jeżeli zachodzi (9.8) to kładziemy $\tau u = u$. Jeżeli nie, to dla $x \in E$, $(\tau u)(x)$ definiujemy jako dowolny element $U(x)$ dla którego

$$\mathcal{P} V_u(x) = q(x, \tau u(x)) + P^{\tau u(x)} V_u(x).$$

Mamy następujący rezultat.

Propozycja 9.1 *Przy założeniu jedyności miary niezmienninczej μ_u dla $u \in V$, mamy*

$$\lambda_u = J((u), X_0) \geq \lambda_{\tau u} = J((\tau u), X_0).$$

Ponadto $\lambda_u > \lambda_{\tau u}$ gdy $\mu_{\tau u}\{x\colon u(x) \neq \tau u(x)\} > 0$.

Dowód Niech $(\lambda_{\tau u}, V_{\tau u})$ będzie rozwiązaniem liniowego równania Bellmana–Howarda dla strategii (τu). Czyli

$$\lambda_{\tau u} + V_{\tau u}(x) = q(x, \tau u(x)) + P^{\tau u(x)} V_{\tau u}(x), \qquad x \in E,$$

$V_{\tau u}(1) = 0$. Wówczas, dla $x \in E$ mamy

$$\begin{aligned} &\lambda_u - \lambda_{\tau u} + V_u(x) - V_{\tau u}(x) \\ &= q(x, u(x)) - q(x, \tau u(x)) + P^{u(x)} V_u(x) - P^{\tau u(x)} V_u(x) \\ &\qquad + P^{\tau u(x)} \left(V_u(x) - V_{\tau u}(x) \right). \end{aligned}$$

Całkując obie względem miary $\mu_{\tau u}$ i korzystając z faktu, że dla miary niezmienniczej $\mu_{\tau u}$ zachodzi

$$\int_E P^{\tau u(x)} \left(V_u(x) - V_{\tau u}(x) \right) \mu_{\tau u}(dx) = \int_E \left(V_u(x) - V_{\tau u}(x) \right) \mu_{\tau u}(dx)$$

otrzymujemy

$$\lambda_u - \lambda_{\tau u} = \int_E f(x) \mu_{\tau u}(dx),$$

gdzie

$$f(x) := q(x, u(x)) - q(x, \tau u(x)) + P^{u(x)} V_u(x) - P^{\tau u(x)} V_{\tau u}(x).$$

Z definicji τu otrzymujemy

$$\begin{aligned} f(x) &= 0, \qquad u(x) = \tau u(x), \\ f(x) &> 0, \qquad u(x) \neq \tau u(x). \end{aligned}$$

Stąd $\lambda_u - \lambda_{\tau u} \geq 0$ i $\lambda_u - \lambda_{\tau u} > 0$ o ile $\mu_{\tau u}\{x \colon u(x) \neq \tau u(x)\} > 0$. $\square$

9.6 Przykłady

9.6.1 Inwestowanie w badania i w reklame

Niech $E = \{0, 1\}$, $U = \{1, 2, 3, 4\}$. Załóżmy, że dane są prawdopodobieństwa przejścia

$$p^1_{0,1} = 0.1,\ p^2_{0,1} = 0.3,\ p^3_{0,1} = 0.7,\ p^4_{0,1} = 0.9,$$

$$p^1_{1,0} = 0.9,\ p^2_{1,0} = 0.8,\ p^3_{1,0} = 0.5,\ p^4_{1,0} = 0.1.$$

Podana jest macierz zysków $q(i, u)$, $i \in E$, $u \in U$,

$$q = \begin{bmatrix} -1 & -2 & -4 & -6 \\ 10 & 8 & 5 & 1 \end{bmatrix}.$$

Celem jest maksymalizacja zysku

$$J((u_n)) = \liminf_{N\to+\infty} \mathbb{E} \sum_{n=0}^{N-1} q(X_n, u_n).$$

Problem możemy zinterpretować następująco: 0 reprezentuje stan firmy przynoszącej straty, a 1 stan przynoszący zyski. Zarząd może inwestować w badania i w reklame zwiększając prawdopodobieństwo przejścia ze stanu 0 w stan 1 oraz zmniejszając prawdopodobieństwo przejścia ze stanu 1 do stanu 0. Parametr $u \in \{1, 2, 3, 4\}$ reprezentuje skale inwestycji.

Rozwiązanie znajdujemy przy pomocy algorytmu Howarda. Przyjmujemy następującą strategie wyjściową: $u(0) = 4 = u(1)$, czyli strategie maksymalnych nakładów. Równania mają postać

$$\begin{aligned} \lambda + V_0 &= q(0,4) + P^u V(0), \\ \lambda + V_1 &= q(1,4) + P^u V(1), \\ V_1 &= 0. \end{aligned}$$

Stąd

$$\begin{aligned} \lambda + V_0 &= -6 + 0.1V_0, \\ \lambda &= 1 + 0.1V_0, \\ V_1 &= 0. \end{aligned}$$

Rozwiązaniem jest układ

$$\lambda = 3/10, \quad V_0 = -7, \quad V_1 = 0.$$

Maksymalizujemy (po i) wyrażenia

$$\begin{aligned} Q(0,i) &:= q(0,i) + \sum_{j=0,1} p^i_{0,j} V_j = q(0,i) + p^i_{0,0} V_0, \\ Q(1,i) &:= q(1,i) + \sum_{j=0,1} p^i_{1,j} V_j = q(1,i) + p^i_{1,0} V_0 \end{aligned}$$

otrzymujemy

$$Q(0,1) = -1 - 7 \times 0.9 = -7.3, \quad Q(0,2) = -2 - 7 \times 0.7 = -6.9,$$

$$Q(0,3) = -4 - 7 \times 0.3 = -6.1, \quad Q(0,4) = -6 - 7 \times 0.1 = -6.7$$

oraz

$$Q(1,1) = 10 - 7 \times 0.9 = 3.7, \quad Q(1,2) = 8 - 7 \times 0.8 = 2.4,$$

$$Q(1,3) = 5 - 7 \times 0.5 = 1.5, \quad Q(1,4) = 1 - 7 \times 0.1 = 0.3.$$

Stąd ulepszoną strategią jest strategia

$$u(0) = 3, u(1) = 1.$$

Dla tej strategii równania mają postać

$$\begin{aligned}\lambda + V_0 &= q(0,3) + p^3_{0.0}V_0,\\ \lambda &= q(1.1) + p^1_{1.0}V_0,\\ V_1 &= 0.\end{aligned}$$

Czyli

$$\begin{aligned}\lambda + V_0 &= -4 + 0.3V_0,\\ \lambda &= 10 + 0.9V_0,\\ V_1 &= 0.\end{aligned}$$

Rozwiązaniem jest układ

$$V_0 = -\frac{35}{4}, \quad V_1 = 0, \quad \lambda = \frac{17}{8}.$$

Maksymalizujemy

$$\begin{aligned}Q(0,i) &= q(0,i) - \frac{35}{4}p^i_{0,0},\\ Q(1,i) &= q(1,i) - \frac{35}{4}p^i_{1,0}.\end{aligned}$$

Mamy

$$\begin{aligned}Q(0,1) &= -1 - \frac{35}{4}\times 0.9 = -\frac{355}{40}, \quad Q(0,2) = -2 - \frac{35}{4}\times 0.7 = -\frac{325}{40},\\ Q(0,3) &= -4 - \frac{35}{4}\times 0.3 = -\frac{265}{40}, \quad Q(0,4) = -6 - \frac{35}{4}\times 0.1 = -\frac{275}{40},\\ Q(1,1) &= 10 - \frac{35}{4}\times 0.9 = \frac{75}{40}, \quad Q(1,2) = 8 - \frac{35}{4}\times 0.8 = \frac{55}{40},\\ Q(1,3) &= 5 - \frac{35}{4}\times 0.5 = \frac{25}{40}, \quad Q(1,4) = 1 - \frac{35}{4}\times 0.1 = \frac{5}{40}.\end{aligned}$$

Stąd strategia $u(0) = 3 = u(1)$ jest optymalna. Ponadto gwarantuje ona średni zysk $\frac{17}{8}$.

9.6.2 Utrzymanie komputera

Ten przykład jest uproszczoną wersją problemu "Kosztów samochodowych," patrz Przykład 1.5 oraz następny paragraf. Zakładamy, że na początku każdego roku podejmujemy decyzje o sprzedaży komputera i kupnie innego lub odkładamy decyzje do przyszłego roku. Przyjmujemy, że interesują nas tylko komputery co najwyżej trzyletnie. Komputer czteroletni musi być sprzedany i zastąpiony nowszym. W poniższej tabeli podajemy odpowiednie ceny kupna i sprzedaży oraz koszty utrzymania. Przyjmujemy, ze komputer dowolnego wieku j może zostać poważnie uszkodzony z prawdopodobienstwem $1 - P_j$, gdzie P_j jest danym prawdopodobieństwami przeżycia jednego roku bez istotnej usterki. Przez poważną usterkę rozumiemy taką, która powoduje, że komputer musi być wymieniony, a jego cena sprzedaży jest równa cenie komputera czteroletniego. Oczywiście naszym celem jest minimalizacja ergodycznego funkcjonału kosztów.

Wiek j	Cena kupna C_j	Cena sprzedaży T_j	Koszty utrzymania K_j	Prawdop. przeżycia P_j
0	10	9	1	0.9
1	8	7	1	0.8
2	6	5	2	0.6
3	4	3	2	0.4
4	2	1	2	0

Pierwsza iteracja Przyjmujemy strategie

$$u(j) = \begin{cases} 0 & \text{dla } j \neq 0 \\ \delta & \text{dla } j = 0. \end{cases}$$

Równania są postaci

$$\begin{aligned} \lambda + V_0 &= K_0 + 0.9V_1, \\ \lambda + V_j &= C_0 - T_j + K_0 + V_0, \qquad j = 1, 2, 3, 4, \\ V_4 &= 0. \end{aligned}$$

Czyli

$$\begin{aligned} \lambda + V_0 &= 1 + 0.9V_1, \\ \lambda + V_j &= 11 - T_j + V_0, \qquad j = 1, 2, 3, 4, \\ V_4 &= 0. \end{aligned}$$

Tak więc

$$\begin{aligned} \lambda + V_0 &= 1 + 0.9V_1, \\ \lambda + V_1 &= 11 - 7 + V_0 = 4 + V_0, \\ \lambda &= 11 - 1 + V_0 = 10 + V_0. \end{aligned}$$

Stąd $V_1 = -6$,

$$2V_0 = -9 + 0.9V_1 = -9 - 5.4 = -14.4, \qquad V_0 - \lambda = -10.$$

A więc rozwiązaniem jest układ

$$V_0 = -7.2, \quad V_1 = -6, \quad V_2 = -4, \quad V_3 = -2, \quad V_4 = 0, \qquad \lambda = 2.8.$$

Minimalizacja Dla danego j szukamy $u(j)$, które minimalizuje wyrażenie

$$Q(j,u) := q(j,u) + P^u V(j). \tag{9.9}$$

Gdy $u \neq \delta$, to

$$q(j,u) + P^u V(j) = C_u - T_j + K_u + V_u. \tag{9.10}$$

Gdy $u = \delta$, wtedy koniecznie $j < 4$ i

$$q(j,\delta) + P^\delta V(j) = K_j + P_j V_{j+1}, \qquad j = 0,1,2,3.$$

Mamy

$$C_u + K_u + V_u = \begin{cases} 10 + 1 - 7.2 = 3.8 & \text{dla } u = 0 \\ 8 + 1 - 6 = 3 & \text{dla } u = 1 \\ 6 + 2 - 4 = 4 & \text{dla } u = 2 \\ 4 + 2 - 2 = 4 & \text{dla } u = 3 \end{cases}$$

Stąd dla wszystkich j optymalne $u \neq \delta$ wynosi 1. Ponaddto

$$Q(j,1) = \begin{cases} 3 - 9 = -6 & \text{dla } j = 0 \\ 3 - 7 = -4 & \text{dla } j = 1 \\ 3 - 5 = -2 & \text{dla } j = 2 \\ 3 - 3 = 0 & \text{dla } j = 3 \\ 3 - 1 = 2 & \text{dla } j = 4 \end{cases}$$

Dla $u = \delta$ mamy

$$Q(j,\delta) = \begin{cases} 1 - 0.9 \times 6 = -4.4 & \text{dla } j = 0 \\ 1 - 0.8 \times 4 = -2.2 & \text{dla } j = 1 \\ 2 - 0.6 \times 2 = 0.8 & \text{dla } j = 2 \\ 2 - 0.4 \times 0 = 2 & \text{dla } j = 3 \end{cases}$$

Stąd ulepszoną strategią jest $u(j) = 1$ dla wszystkich j.
Druga iteracja Dla strategi $u(j) = 1$, $j = 0,1,2,3,4$ mamy następujące równania

$$\lambda + V_j = C_1 - T_j + K_1 + V_1, \quad j = 0,1,2,3,4, \qquad V_4 = 0.$$

Stąd

$$\lambda = C_1 - T_1 + K_1 = 2, \qquad V_1 = \lambda - C_1 + T_4 - K_1 = 2 - 8 + 1 - 1 = -6.$$

Tak więc rozwiązaniem jest układ

$$V_0 = -8, \quad V_1 = -6, \quad V_2 = -4, \quad V_3 = -2, \quad V_4 = 0, \qquad \lambda = 2.$$

Minimalizacja Jak w pierwszej iteracji szukamy strategii u minimalizującej (9.9) z nowymi (V_j). Z (9.10) dla $u \neq \delta$ mamy

$$Q(j,u) = \begin{cases} 11 + V_0 - T_j = 3 - T_j & \text{dla } u = 0 \\ 9 + V_1 - T_j = 3 - T_j & \text{dla } u = 1 \\ 8 + V_2 - T_j = 4 - T_j & \text{dla } u = 2 \\ 6 + V_3 - T_j = 4 - T_j & \text{dla } u = 3. \end{cases}$$

Dla $u = \delta$ mamy

$$Q(j,\delta) = \begin{cases} 1 - 0.9 \times 6 = -4.4 & \text{dla } j = 0 \\ 1 - 0.8 \times 4 = -2.2 & \text{dla } j = 1 \\ 2 - 0.6 \times 2 = 0.8 & \text{dla } j = 2 \\ 2 - 0.4 \times 0 = 2 & \text{dla } j = 3. \end{cases}$$

Stąd strategią $u(j) = 1$ dla wszystkich j minimalizuje Q, a więc jest optymalna dla naszego problemu. Ponadto optymalny koszt λ wynosi 2.

9.6.3 Koszty samochodowe

Za [11] podamy rozwiązanie problemu przedstawionego w Przykładzie 1.5. Dane (w dolarach) podane są w poniższych tabelach.

Wiek j	Cena kupna C_j	Cena sprzedaży T_j	Koszty utrzymania K_j	Prawdop. przeżycia P_j
0	2000	1600	50	1
1	1840	1460	53	0.999
2	1680	1340	56	0.0998
3	1560	1230	59	0.997
4	1300	1050	62	0.996
5	1220	980	65	0.994
6	1150	910	68	0.991
7	1080	840	71	0.988
8	900	710	75	0.985
9	840	650	78	0.983
10	900	710	75	0.985
11	730	550	84	0.975
12	600	480	87	0.970
13	560	430	90	0.965
14	520	390	93	0.960
15	480	360	96	0.955
16	440	330	100	0.950
17	420	310	103	0.945
18	400	290	106	0.940
19	380	270	109	0.935
20	360	255	112	0.930

Wiek j	Cena kupna C_j	Cena sprzedaży T_j	Koszty utrzymania K_j	Prawdop. przeżycia P_j
21	345	240	115	0.925
22	330	225	118	0.919
23	315	210	121	0.910
24	300	200	125	0.900
25	290	190	129	0.890
26	280	180	133	0.880
27	265	170	137	0.865
28	250	160	141	0.850
29	240	150	145	0.820
30	230	145	150	0.790
31	220	140	155	0.760
32	210	135	160	0.730
33	200	130	167	0.660
34	190	120	175	0.590
35	180	115	182	0.510
36	170	110	190	0.430
37	160	105	205	0.300
38	150	95	220	0.200
39	140	87	235	0.100
40	130	80	250	0

W [11] przyjęto następującą strategie wyjściową

$$u(j) = \begin{cases} 36 & \text{gdy } j \leq 20, \\ \delta & \text{gdy } 20 < j < 40, \\ 36 & \text{gdy } j = 40. \end{cases}$$

Rozwiązanie zostało znalezione po siedmiu iteracjach. Otrzymano następującą strategie optymalną: jeżeli samochód na lat nie mniej niż $\frac{1}{2}$ a nie więcej niż $6\frac{1}{2}$, to należy go zatrzymać. W innym przypadku należy go sprzedać i kupić 3 letni. Oczywiście postępując zgodnie z tą strategią dochodzimy do następującego cyklu. Zachowujemy samochód mający od 3 do $6\frac{1}{2}$ lat a następnie sprzedamy go i kupimy 3 letni. Koszt przy wyjściowej strategii wynosił 250\$. Optymalny koszt wynosi 150.95\$.

Część I

Filtracja

10

Wstęp

Niech $Z(t), Y(t)$, $t \geq 0$ będą procesami stochastycznymi o wartościach w przestrzeniach mierzalnych $(E_1, \mathcal{E}_1)$, $(E_2, \mathcal{E}_2)$. Zagadnienie filtracji polega na wyznaczeniu, w dowolnej chwili t, najlepszego estymatora zmiennej $Z(t)$ w oparciu o obserwacje $Y(s)$, $s \leq t$. Proces Y nazywa się *procesem obserwacji*. Rozważmy na przykład sytuację

$$\begin{aligned} dZ(t) &= f(Z(t))dt + dW_1(t), \quad Z(0) = z_0 \in \mathbb{R}^k \\ dY(t) &= h(Z(t))dt + dW_2(t), \quad Y(0) = y_0 \in \mathbb{R}^l, \end{aligned}$$

w której W_1 i W_2 są niezależnymi procesami Wienera. Przy założeniu, że $f\colon \mathbb{R}^k \to \mathbb{R}^k$ i $h\colon \mathbb{R}^l \to \mathbb{R}^l$ spełniają warunek Lipschitza układ ten ma dokładnie jedno rozwiązanie. Proces Y ewoluuję, na ogół, w przestrzeni o mniejszym wymiarze niż proces częściowo obserwowalny proces Z. Gdyby obserwacja nie była zaburzona, to

$$Y(t) = y_0 + \int_0^t h(Z(s))ds, \qquad t \geq 0$$

i nawet w tej sytuacji, ze znajomości trajektorii Y nie zawsze dałoby się odtworzyć dokładnie stanu Z.

11

Estymatory i warunkowa wartość oczekiwana

Niech $(\Omega, \mathcal{F}, \mathbb{P})$ będzie ustaloną przestrzenią probabilistyczną. Aby zagadnienie filtracji ściśle sformułować rozpatrzmy najpierw klasyczne zagadnienie estymacji wartości zmiennej losowej Z, przyjmującej wartości rzeczywiste, gdy obserwowana jest wartość zmiennej losowej Y o wartościach w $(E, \mathcal{E})$. *Estymatorem* nazywa się dowolną funkcję mierzalną ψ z przestrzeni E w zbiór liczb rzeczywistych. Estymator $\widehat{\psi}$ jest *optymalny*, gdy dla dowolnego estymatora ψ,

$$\mathbb{E}\left(\widehat{\psi}(Y) - Z\right)^2 \leq \mathbb{E}\left(\psi(Y) - Z\right)^2.$$

Estymator, jest więc obiektem deterministycznym, a wartość zmiennej losowej $\psi(Y)$ to estymator zmiennej Z przypisywany obserwacji Y. Liczbę

$$\mathbb{E}\left(\psi(Y) - Z\right)^2$$

interpretujemy jako średni błąd kwadratowy popełniany, gdy stosuje się estymator ψ. Naszym pierwszym celem będzie udowodnienie istnienia optymalnego estymatora i wyznaczenie jego postaci.

Propozycja 11.1 *Jeżeli* $\mathbb{E}\, Z^2 < +\infty$*, to stnieje optymalny estymator* $\widehat{\psi}$ *zmiennej* Z *ze względu na* Y*. Jest on dowolną funkcją mierzalną, taką że*

$$\widehat{\psi}(Y) = \mathbb{E}\left(Z|\sigma(Y)\right). \tag{11.1}$$

Dowód Zauważmy, że funkcja $\widehat{\psi}$ dla której zachodzi (11.1) zawsze istnieje na podstawie Lematu 1.1. Połóżmy, $\mathcal{G} = \sigma(Y)$. Przypomnijmy, patrz Propozycja 1.1, że $\mathbb{E}\left(Z|\mathcal{G}\right)$ jest rzutem ortogonalny zmiennej losowej Z na podprzestrzeń $L^2(\Omega, \mathcal{G}, \mathbb{P})$ przestrzeni $L^2(\Omega, \mathcal{F}, \mathbb{P})$. Iloczyn skalarny na $L^2(\Omega, \mathcal{G}, \mathbb{P})$ dany jest wzorem

$$\langle X_1, X_2 \rangle = \mathbb{E}\, X_1 X_2.$$

Niech $||\cdot||$ będzie normą na $L^2(\Omega, \mathcal{G}, \mathbb{P})$. Z definicji rzutu wynika, że

$$||Z - Z'|| \geq ||Z - \mathbb{E}\left(Z|\mathcal{G}\right)||, \qquad \forall\, Z' \in L^2(\Omega, \mathcal{G}, \mathbb{P}).$$

Jeśli więc $\mathbb{E}\,(\psi(Y))^2 < \infty$, to

$$\mathbb{E}\,||Z - Z'||^2 \geq \mathbb{E}\,||Z - \mathbb{E}\,(Z|\mathcal{G})||^2, \qquad \forall\, Z' \in L^2(\Omega, \mathcal{G}, \mathbb{P}). \tag{11.2}$$

Jeśli $\mathbb{E}\,(\psi(Y))^2 = +\infty$, to $\mathbb{E}\,(Z - \psi(Y))^2 = +\infty$, czyli (11.2) zachodzi i w tym przypadku. $\square$

Odtąd, optymalny estymator utożsamiać będziemy przez (11.1) z warunkową wartością oczekiwaną, również gdy $\mathbb{E}\,Z^2 = +\infty$.

12

Filtr Kalmana–Bucy

12.1 Warunkowanie zmiennych gaussowskich

Niech $m \in \mathbb{R}^n$, $Q \in M(n \times n)$. Przez $\mathcal{N}(m, Q)$ oznaczamy *rozkład gaussowski* w $\mathbb{R}^n$ ze średnią m i macierzą kowariancji Q. Gdy Q jest ściśle dodatnio określona to rozkład $\mathcal{N}(m, Q)$ ma gęstość

$$g(x) = \frac{1}{\sqrt{(2\pi)^n \det Q}} \exp\left\{-\frac{1}{2}\left\langle Q^{-1}(x-m),\, x-m\right\rangle\right\}, \qquad x \in \mathbb{R}^n.$$

W teorii filtracji zasadnicze znaczenie mają wzory na rozkłady warunkowe w przypadku gdy wielowymiarowa zmienna losowa (X, Y) jest gaussowska.

Twierdzenie 12.1 *Załóżmy, że gaussowska zmienna losowa (X, Y) przyjmująca wartości w $\mathbb{R}^k \times \mathbb{R}^l$ ma wektor wartości oczekiwanych (m_X, m_Y) i macierz kowariancji*

$$Q = \begin{pmatrix} Q_{XX}, \, Q_{XY} \\ Q_{YX}, \, Q_{YY} \end{pmatrix}.$$

Wówczas rozkład warunkowy $\mathbb{P}\,(X|Y)$ zmiennej X względem Y jest gaussowski. Ponadto

$$\mathbb{P}\,(X|Y)(\omega) = \mathcal{N}(\widehat{m}(Y(\omega)), \widehat{Q}), \qquad \omega \in \Omega,$$

gdzie

$$\begin{aligned} \widehat{m}(y) &= m_X + Q_{XY} Q_{YY}^{-1}(y - m_y), \\ \widehat{Q} &= Q_{XX} - Q_{XY} Q_{YY}^{-1} Q_{YX}. \end{aligned}$$

Dowód Wyliczymy warunkową gęstość przy założeniu, że wektor gaussowski $(X, Y) \in \mathbb{R}^k \times \mathbb{R}^l$ ma gęstość

$$\begin{aligned} &g(x, y) \\ &= \frac{1}{\sqrt{(2\pi)^{k+l} \det Q}} \exp\left\{-\frac{1}{2}\left\langle Q^{-1}\begin{pmatrix} x - m_X \\ y - m_Y \end{pmatrix}, \begin{pmatrix} x - m_X \\ y - m_Y \end{pmatrix}\right\rangle\right\} \end{aligned}$$

Niech

$$Q = \begin{pmatrix} Q_{XX} \ , Q_{XY} \\ Q_{YX} \ , Q_{YY} \end{pmatrix}, \qquad Q^{-1} = R = \begin{pmatrix} R_{11} \ , R_{12} \\ R_{21} \ , R_{22} \end{pmatrix}.$$

Z Lematu 1.3 mamy

$$\begin{aligned} g(x|y) &= c_1(y) \exp -\frac{1}{2} \{\langle R_{11}(x-m_X) + R_{12}(y-m_Y), x - m_X \rangle \\ &\qquad + \langle R_{21}(x-m_X) + R_{22}(y-m_Y), y-m_Y \rangle\} \\ &= c_2(y) \exp -\frac{1}{2} \left\{ \left\langle R_{11} \left((x-m_X) + R_{11}^{-1} R_{12}(y-m_Y) \right), \right.\right. \\ &\qquad \left.\left. x - m_X + R_{11}^{-1} R_{12}(y-m_Y) \right\rangle \right\}, \end{aligned}$$

bo

$$\langle R_{11}(x-m_X), R_{11}^{-1} R_{12}(y-m_Y) \rangle = \langle R_{21}(x-m_X), y-m_Y \rangle.$$

Stąd

$$\begin{aligned} g(x|y) = c_2(y) \exp -\frac{1}{2} & \left\langle R_{11} \left(x - \left[m_X - R_{11}^{-1} R_{12}(y-m_Y) \right] \right), \right. \\ & \left. x - \left[m_X - R_{11}^{-1} R_{12}(y-m_Y) \right] \right\rangle. \end{aligned}$$

Stąd wynika, że gęstość warunkowa jest gaussowska z wektorem wartości średniej

$$\widehat{m}(y) = M_X - R_{11}^{-1} R_{12}(y-m_Y)$$

oraz z macierzą kowariancji R_{11}^{-1}.

Ponieważ $QR = I$ mamy

$$Q_{XX} R_{11} + Q_{XY} R_{21} = I, \qquad Q_{YX} R_{11} + Q_{YY} R_{21} = 0.$$

Dlatego

$$\begin{aligned} R_{11}^{-1} &= Q_{XX} + Q_{XY} R_{21} R_{11}^{-1}, \\ Q_{YX} = -Q_{YY} R_{21} R_{11}^{-1}, & \qquad Q_{YY}^{-1} Q_{YX} = -R_{21} R_{11}^{-1}. \end{aligned}$$

Czyli

$$\begin{aligned} R_{11}^{-1} &= Q_{XX} - Q_{XY} Q_{YY}^{-1} Q_{YX}, \\ R_{11}^{-1} R_{12} &= -Q_{XY} Q_{YY}^{-1}. \end{aligned}$$

Ostatecznie

$$\begin{aligned} \widehat{m}(y) &= m_X + Q_{XY} Q_{YY}^{-1} (y - m_Y), \\ \widehat{Q} &= Q_{XX} - Q_{XY} Q_{YY}^{-1} Q_{YX}. \end{aligned}$$

□

Wniosek 12.1 *Warunkowa wartość oczekiwana* $\mathbb{E}(X|Y)$ *jest identyczna z rzutem ortogonalnym* X *na podprzestrzeń zmiennych losowych*

$$\left\{ a\chi_{\mathbb{R}^l} + BY \colon a \in \mathbb{R}^k,\ B \in M(l \times k) \right\} := L_k^2(\chi, Y) \subset L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^k).$$

Wniosek 12.2 *Jeżeli* (X, Y, Z) *jest gaussowskim wektorem losowym, takim że* Y *i* Z *są niezależne, to*

$$\mathbb{E}\,(X|Y,Z) = \mathbb{E}\,(X|Y) + \mathbb{E}\,(X|Z) - \mathbb{E}\,X.$$

Istotnie. Możemy założyć, że $\mathbb{E}\,Y = 0$ i $\mathbb{E}\,Z = 0$. Z pierwszego wniosku wynika, że $\mathbb{E}\,(X|Y,Z)$ jest rzutem ortogonalnym na przestrzeń

$$L^2_k(\chi, Y, Z)$$

będącą sumą dwóch przestrzeni ortogonalnych

$$L^2_k(\chi, Y) \qquad \text{i} \qquad L^2_k(\chi, Z).$$

Wystarczy więc zauważyć, że rzut ortogonalny na przestrzeń $L^2_k(\chi, Z)$ jest równy $\mathbb{E}\,(X|Z) - \mathbb{E}\,X$. $\square$

12.2 Zasadniczy rezultat

Zajmiemy się problemem filtracji dla układu liniowego

$$X_{n+1} = \Psi X_n + \xi_{n+1}, \qquad Y_n = \Theta X_n + \eta_n, \qquad n = 0, 1, \ldots.$$

O zmiennych losowych $\{\xi_n\}$ zakładamy, że są niezależne, gaussowskie o wartościach w $\mathbb{R}^d$. Mają one średnią zero i macierz kowariancji R_1. Podobnie zmienne $\{\eta_n\}$ są niezależne, gaussowskie, o wartościach w $\mathbb{R}^m$ z wartością oczekiwaną zero i macierzą kowariancji R_2. Dodatkowo zakładamy, że $\{\xi_n\}$ są niezależne od $\{\eta_n\}$. Zmienna X_0 jest gaussowska i niezależna od $\{\xi_n\}$ i $\{\eta_n\}$. Ma wartość oczekiwaną $\mathbb{E}\,X_0$ i kowariancje R_0.

Przez (S_n) oznaczamy ciąg macierzy zdefiniowany rekurencyjnie

$$\begin{aligned} S_{n+1} &= \Psi S_n \left(I + \Theta^* R_2^{-1} \Theta S_n\right)^{-1} \Psi^* + R_1, \qquad n = 0, 1, \ldots, \\ S_0 &= R_0. \end{aligned}$$

Niech

$$\widehat{X}_n = \mathbb{E}\,(X_n|Y_0, \ldots, Y_n)\,, \qquad n = 0, 1, \ldots.$$

Zachodzi następujące twierdzenie Kalmana–Bucy.

Twierdzenie 12.2 *Załóżmy, że macierz* R_2 *jest odwracalna. Niech*

$$K_n = S_n \Theta^* \left(R_2 + \Theta S_n \Theta^*\right)^{-1}, \qquad n = 0, 1, \ldots.$$

Wtedy

$$\begin{aligned} \widehat{X}_{n+1} &= \Psi \widehat{X}_n + K_{n+1}\left(Y_{n+1} - \Theta \Psi \widehat{X}_n\right), \qquad n = 0, 1, \ldots, \\ \widehat{X}_0 &= \mathbb{E}\,\widehat{X}_0 + K_0\left(Y_0 - \mathbb{E}\,Y_0\right). \end{aligned}$$

Ponadto zmienne losowe

$$\begin{aligned} Z_n &= Y_n - \Theta\Psi\widehat{X}_{n-1}, \qquad n = 1, 2. \ldots, \\ Z_0 &= Y_0 - \mathbb{E}\, Y_0 \end{aligned}$$

są niezależne i gaussowskie. Ciąg $(\widehat{X}_n)$ *jest gaussowskim łańcuchem Markowa. Rozkłady warunkowe* X_n *przy zadanych* $Y_0, \ldots Y_n$ *są gaussowskie z wartością oczekiwaną* $\widehat{X}_n$ *i kowariancją*

$$P_n = S_n \left(I + \Theta^* R_2 \Theta S_n \right)^{-1}, \qquad n = 0, 1, \ldots.$$

Dowód Dla prostoty będziemy przez $[X, Y]$ oznaczać macierz kowariancji X i Y. Ponadto kładziemy $[X] = [X, X]$. Tak więc dla wektora gaussowskiego (X, Y) mamy

$$\widehat{X} = \mathbb{E}\,(X|Y) = \mathbb{E}\, X + [X, Y][Y]^{-1}(Y - \mathbb{E}\, Y)$$

oraz

$$\begin{aligned} \widehat{Q} &= \mathbb{E}\left((X - \widehat{X}) \otimes (X - \widehat{X}) | Y \right) = \mathbb{E}\,(X - \widehat{X}) \otimes (X - \widehat{X}) \\ &= [X] - [X, Y]\,[Y]^{-1}[Y, X]. \end{aligned}$$

Będziemy korzystali z faktu, że gdy zmienne losowe gaussowskie X i Y są niezależne, to

$$[X, Y] = 0 = [X - \widehat{X}, \widehat{X}].$$

Przechodzimy teraz do dowodu twierdzenia. Przyjmijmy oznaczenia

$$\tilde{X}_n = \mathbb{E}\,(X_n | Y_0, \ldots, Y_{n-1}), \qquad n = 1, 2, \ldots$$

oraz $\tilde{X}_0 = \mathbb{E}\, X_0$. Ponadto niech

$$\begin{aligned} \overline{P}_n &= \mathbb{E}\,(X_n - \widehat{X}_n) \otimes (X_n - \widehat{X}_n), \\ \overline{S}_n &= \mathbb{E}\,(X_n - \tilde{X}_n) \otimes (X_n - \tilde{X}_n). \end{aligned}$$

Kluczową role w dowodzie odgrywa *ciąg innowacyjny*

$$Z_0 = Y_0 - \mathbb{E}\, Y_0, \qquad Z_{n+1} = Y_{n+1} - \mathbb{E}\,(Y_{n+1} | Y_0, \ldots, Y_n).$$

Zmienne losowe $\{Z_n\}$ są gaussowskie, niezależne i o średniej 0. Indukcjnie łatwo pokazuje się, że dla $n = 0, 1, \ldots$ zachodzi

$$\begin{aligned} &\left\{ a + bY_0 + \ldots B_n Y_n \colon a \in \mathbb{R}^k,\ B_i \in M(k \times l) \right\} \\ &= \left\{ a + bZ_0 + \ldots B_n Z_n \colon a \in \mathbb{R}^k,\ B_i \in M(k \times l) \right\}. \end{aligned}$$

By wyprowadzić równanie na $\widehat{X}_n$ zauważmy, że ponieważ ξ_{n+1} jest niezależne od $\mathcal{Y}_n$,

$$\mathbb{E}\,(\xi_{n+1}|Z_0,\ldots,Z_n)=\mathbb{E}\,\xi_{n+1}=0.$$

Stąd

$$\begin{aligned}\widehat{X}_{n+1}&=\mathbb{E}\,(X_{n+1}|Y_0,\ldots,Y_{n+1})=\mathbb{E}\,(X_{n+1}|Z_0,\ldots,Z_{n+1})\\&=\mathbb{E}\,(X_{n+1}|Z_0,\ldots Z_n)+\mathbb{E}\,(X_{n+1}|Z_{n+1})-\mathbb{E}\,X_{n+1}\\&=\mathbb{E}\,(\Psi X_n+\xi_{n+1}|Z_0,\ldots Z_n)+\mathbb{E}\,X_{n+1}\\&\qquad+[X_{n+1},Z_{n+1}][Z_{n+1}]^{-1}Z_{n+1}-\mathbb{E}\,X_{n+1}\\&=\Psi\widehat{X}_n+[X_{n+1},Z_{n+1}][Z_{n+1}]^{-1}Z_{n+1}.\end{aligned}$$

Co więcej

$$\begin{aligned}Z_{n+1}&=Y_{n+1}-\mathbb{E}\,(\Theta\Psi X_n+\Theta\xi_{n+1}|Y_0,\ldots,Y_n)\\&=Y_{n+1}-\Theta\Psi\widehat{X}_n.\end{aligned}$$

Wykażemy teraz równość

$$[X_n,Z_n][Z_n]^{-1}=K_n,\qquad n=0,1,\ldots.$$

W ten sposób dowód pierwszej części twierdzenia będzie zakończony. Ponieważ

$$\begin{aligned}Z_n&=\Theta X_n+\eta_n-\mathbb{E}\,(\Theta X_n+\eta_n|Y_0,\ldots,Y_{n-1})\\&=\Theta(X_n-\tilde{X}_n)+\eta_n,\end{aligned}$$

dlatego

$$[Z_n]=\Theta[X_n-\tilde{X}_n]\Theta^*+R_2=\Theta\overline{S}_n\Theta^*+R_2,$$

gdzie z definicji

$$\overline{S}_n=[X_n-\tilde{X}_n].$$

Podobnie

$$\begin{aligned}[X_n,Z_n]&=\Big[X_n,\Theta(X_n-\tilde{X}_n)+\eta_n\Big]=[X_n,X_n-\tilde{X}_n]\Theta^*\\&=[X_n-\tilde{X}_n]\Theta^*=\overline{S}_n\Theta^*.\end{aligned}$$

Stąd

$$K_n=[X_n,Z_n][Z_n]^{-1}=\overline{S}_n\Theta^*\left(\Theta\overline{S}_n\Theta^*+R_2\right)^{-1}.$$

Wyprowadzimy teraz wzory rekurencyjne na $\overline{S}_n$ i $\overline{P}_n$. Otóż mamy

$$\begin{aligned}X_n-\tilde{X}_n&=X_n-\mathbb{E}\,(X_n|Y_0,\ldots,Y_{n-1})\\&=\Psi X_{n-1}+\xi_n-\mathbb{E}\,(\Psi X_{n-1}+\xi_n|Y_0,\ldots,Y_{n-1})\\&=\Psi\left(X_{n-1}-\widehat{X}_{n-1}\right)+\xi_n.\end{aligned}$$

Dlatego

$$\overline{S}_n=\Psi\overline{P}_{n-1}\Psi^*+R_1.$$

Co więcej

$$\begin{aligned}
\overline{P}_n &= [X_n - \widehat{X}_n] = [X_n - \mathbb{E}\,(X_n|Z_0, \ldots, Z_n)] \\
&= \Big[(X_n - \tilde{X}_n) - K_n Z_n\Big] \\
&= [X_n - \tilde{X}_n] - [X_n - \tilde{X}_n, Z_n]K_n^* - K_n[Z_n, X_n - \tilde{X}_n] \\
&\qquad +K_n[Z_n]K_n^* \\
&= [X_n - \tilde{X}_n] - [X_n, Z_n]K_n^* - K_n[Z_n, X_n] + K_n[Z_n]K_n^* \\
&= [X_n - \tilde{X}_n] - K_n[Z_n]K_n^* - K_n[Z_n]K_n^* + K_n[Z_n]K_n^* \\
&= [X_n - \tilde{X}_n] - K_n[Z_n]K_n^* \\
&= \overline{S}_n - \overline{S}_n \Theta^* K_n^* \\
&= \overline{S}_n - \overline{S}_n \Theta^* \left(\Psi^* \overline{S}_n \Theta + R_2\right)^{-1} \Theta \overline{S}_n \\
&= \overline{S}_n \left(I + \Theta^* R_2^{-1} \Theta \overline{S}_n\right)^{-1}
\end{aligned}$$

skąd wynikają poszukiwane wzory. □

13

Sterowanie z niepełną informacją

Załóżmy, że proces $\{X_n\}$ zadany jest rekurencyjnie

$$X_{n+1} = \Psi X_n + \Phi u_n + \xi_n,$$

gdzie $\Psi \in M(d \times d)$, $\Phi \in M(d \times l)$, a $\{\xi_n\}$ jest ciągiem niezależnych gaussowskich zmiennych losowych w $\mathbb{R}^d$ o średnich zero i kowariancji R_1. Tak jak w Rozdziale 2 funkcjonał kosztu (lub zysku) dany jest wzorem

$$J_N(\pi, X_0) = \mathbb{E}\left(\sum_{n=0}^{N-1} q_n(X_n, u_n) + r_N(X_N)\right),$$

gdzie $N < +\infty$ jest skończonym horyzontem czasowym, a $\{q_n\}$ i r_N są mierzalnymi nieujemnymi funkcjami określonymi odpowiednio na $\mathbb{R}^d \times \mathbb{R}^l$ i $\mathbb{R}^d$. Obserwujemy jednak nie proces X ale proces Y dany wzorem

$$Y_n = \Theta X_n + \eta_{n+1},$$

gdzie $\Theta \in M(m \times d)$ a $\{\eta_n\}$ jest ciągiem niezależnych zmiennych gaussowskich w $\mathbb{R}^m$ o średniej zero i macierzy kowariancji R_2. Tak więc wymagamy, aby strategie były funkcjami Y a nie X. Dokładnie strategia

$$\pi = (u_0, \ldots, u_{N-1})$$

jest ciągiem funkcji mierzalnych

$$u_n \colon \mathbb{R}^m \times \ldots \mathbb{R}^m = \mathbb{R}^{m \times (n+1)} \to \mathbb{R}^l,$$

a proces jej odpowiadający (X_n^π) dany jest wzorem

$$\begin{cases} X_{n+1}^\pi = \Psi X_n^\pi + \Phi u_n(Y_0^\pi, \ldots, Y_n^\pi) + \xi_{n+1}, \\ Y_n^\pi \quad = \Theta X_n^\pi + \eta_n. \end{cases}$$

Gdy J_N jest kosztem to szukamy strategii $\widehat{\pi}$ która go minimalizuje, gdy J_N jest zyskiem to szukamy $\widehat{\pi}$, która go maksymalizuje.

Naszym celem będzie sprowadzenie problemu częściowo obserwowanego do problemu sterowania przy pełnej obserwacji. Niech

$$\pi = (u_0, \ldots, u_{N-1}) = \big(u_0(Y_0^\pi), \ldots, u_{N-1}(Y_0^\pi, \ldots, Y_{N-1}^\pi)\big)$$

będzie sterowaniem dopuszczalnym dla naszego problemu z niepełną informacją. Zdefiniujmy procesy

$$\{(\overline{X}_n, \overline{Y}_n)\} \quad \text{i} \quad \{(\tilde{X}_n, \tilde{Y}_n)\}$$

w następujący sposób

$$\overline{X}_0 = X_0, \qquad \overline{X}_{n+1} = \Psi \overline{X}_n + \xi_{n+1}, \qquad \overline{Y}_n = \Theta \overline{X}_n + \eta_n$$

oraz

$$\tilde{X}_0 = 0, \qquad \tilde{X}_{n+1} = \Psi \tilde{X}_n + \Phi u_n, \qquad \tilde{Y}_n = \Theta \tilde{X}_n.$$

Wówczas

$$X_n = \overline{X}_n + \tilde{X}_n, \qquad Y_n = \overline{Y}_n + \tilde{Y}_n.$$

Niech

$$\mathcal{Y}_n = \sigma(Y_0, \ldots, Y_n), \qquad \overline{\mathcal{Y}}_n = \sigma(\overline{Y}_0, \ldots, \overline{Y}_n).$$

Lemat 13.1 *Dla dowolnego n mamy*

$$\mathcal{Y}_n = \overline{\mathcal{Y}}_n.$$

Ponadto $\tilde{X}_{n+1}$ jest $\mathcal{Y}_n$ mierzalne.

Dowód Teza lematu jest prawdziwa gdy $n = 0$ bo $\tilde{X}_0 = 0$, a więc

$$Y_0 = \Theta X_0 + \eta_0 = \Theta \overline{X}_0 + \eta_0 = \overline{Y}_0$$

oraz

$$X_1 - \overline{X}_1 = \Psi(X_0 - \overline{X}_0) + \Phi u_0 = \Phi u_0.$$

Załóżmy, że

$$\mathcal{Y}_0 = \overline{\mathcal{Y}}_0, \ldots \mathcal{Y}_n = \overline{\mathcal{Y}}_n$$

oraz, że $X_j - \overline{X}_j$ jest $\mathcal{Y}_{j-1}$-mierzalne dla $j \leq n$. Ponieważ

$$X_{n+1} - \overline{X}_{n+1} = \Psi(X_n - \overline{X}_n) + \Phi u_n,$$

zmienna $X_{n+1} - \overline{X}_{n+1}$ jest $\mathcal{Y}_n = \overline{\mathcal{Y}}_n$-mierzalna. Ale

$$Y_{n+1} - \overline{Y}_{n+1} = \Theta(X_{n+1} - \overline{X}_{n+1}).$$

Stąd Y_{n+1} jest $\overline{\mathcal{Y}}_{n+1}$-mierzalna i $\overline{Y}_{n+1}$ jest $\mathcal{Y}_{n+1}$-mierzalny, co kończy dowód lematu. □

Niech $\mathbb{P}(\overline{X}_n|\mathcal{Y}_n)$, $n = 0, 1, \ldots$ będzie rozkładem warunkowym $\overline{X}_n$ względem $\overline{Y}_0, \ldots, \overline{Y}_n$. Z twierdzenia Kalmana–Bucy,

$$\mathbb{P}\,(\overline{X}_n|\mathcal{Y}_n) = \mathcal{N}(\widehat{\overline{X}}_n, P_n),$$

gdzie $\{P_n\}$ jest ciągiem macierzy kowariancji występujący w twierdzeniu, a

$$\widehat{\overline{X}}_n = \mathbb{E}\,(\overline{X}_n|\overline{Y}_0,\ldots,\overline{Y}_n) = \mathbb{E}\,(\overline{X}_n|\overline{\mathcal{Y}}_n).$$

Kluczowym faktem dla sprowadzenia problemu z częściową obserwacją do problemu z pełną obserwacją jest następujące twierdzenie.

Twierdzenie 13.1 *Dla każdego sterowania dopuszczalnego* π *rozkład warunkowy* $\mathbb{P}\,(X_n|\mathcal{Y}_n)$ *zmiennej* X_n *względem* $(Y_0,\ldots,Y_n)$ *jest gaussowski. Ponadto*

$$\mathbb{E}\,(X_n|\mathcal{Y}_n) = \widehat{\overline{X}_n} + \tilde{X}_n$$

oraz

$$\mathbb{P}\,(X_n|\mathcal{Y}_n) = \mathcal{N}(\widehat{\overline{X}_n} + \tilde{X}_n, P_n).$$

Dowód Ponieważ $X_n = X_n + \tilde{X}_n$ oraz $\mathcal{Y}_n = \overline{\mathcal{Y}}_n$, oraz ponieważ $\tilde{X}_n$ jest $\overline{\mathcal{Y}}_{n-1}$-mierzalna, a więc i $\overline{\mathcal{Y}}_n$-mierzalna, mamy

$$\mathbb{E}\,(X_n|\mathcal{Y}_n) = \widehat{\overline{X}_n} + \tilde{X}_n$$

oraz

$$\begin{aligned}
\mathbb{E}\,(\psi(X_n)|\mathcal{Y}) &= \mathbb{E}\,(\psi(\overline{X}_n + \tilde{X}_n)|\overline{\mathcal{Y}}_n)\\
&= \int_{\mathbb{R}^d} \psi(x + \tilde{X}_n)\mathcal{N}(\widehat{\overline{X}_n}, P_n)(dx)\\
&= \int_{\mathbb{R}^d} \psi(x)\mathcal{N}(\widehat{\overline{X}_n} + \tilde{X}_n, P_n)(dx),
\end{aligned}$$

co dowodzi twierdzenia. □

Wyliczmy teraz ewolucję $\widehat{X}_n =: \widehat{\overline{X}}_n + \tilde{X}_n$. Pokażemy, że estymator $\widehat{X}_n$ można wyliczyć rekurencyjnie w terminach procesu obserwacji (Y_n). Pozwala to rozwiązywać zagadnienie "on line". Mamy następujący rezultat.

Propozycja 13.1 *Dla* $n = 0, 1, \ldots$ *mamy*

$$\widehat{X}_{n+1} = \Psi\widehat{X}_n + \Phi u_n + K_{n+1}\overline{Z}_{n+1},$$

gdzie $\{\overline{Z}_n\}$ *jest ciągiem niezależnych gaussowskich zmiennych losowych o zerowych średnich i macierzy kowariancji*

$$[\overline{Z}_n] = \Theta S_n \Theta^* + R_2, \qquad n = 0, 1, \ldots.$$

Ponadto

$$\widehat{X}_{n+1} = (I - K_{n+1}\Theta)\Psi\widehat{X}_n + (I - K_{n+1}\Theta)\Phi u_n + K_{n+1}Y_{n+1}.$$

Dowód Przy oznaczeniach z twierdzenia Kalmana–Bucy zachodzi

$$\begin{aligned}
\widehat{X}_{n+1} &= \widehat{\overline{X}}_{n+1} + \tilde{X}_{n+1} \\
&= \Psi \widehat{\overline{X}}_n + K_{n+1}\overline{Z}_{n+1} + \tilde{X}_{n+1} \\
&= \Psi \widehat{\overline{X}}_n + \Psi \tilde{X}_n + \Phi u_n + K_{n+1}\overline{Z}_{n+1} \\
&= \Psi \widehat{X}_n + \Phi u_n + K_{n+1}\overline{Z}_{n+1},
\end{aligned}$$

co dowodzi pierwszej części. Następnie zachodzi

$$\begin{aligned}
\overline{Z}_{n+1} &= \overline{Y}_{n+1} - \Theta\Psi \widehat{\overline{X}}_n \\
&= \overline{Y}_{n+1} - \Theta\Psi \left(\widehat{X}_n - \tilde{X}_n\right) \\
&= Y_{n+1} + \Theta\left(\overline{X}_{n+1} - X_{n+1}\right) - \Theta\Psi\left(\widehat{X}_n - \tilde{X}_n\right) \\
&= Y_{n+1} + \Theta\left(\Psi(\overline{X}_n - X_n) - \Phi u_n\right) - \Theta\Psi\left(\widehat{X}_n - \tilde{X}_n\right) \\
&= Y_{n+1} - \Theta\Psi\widehat{X}_n - \Theta\Phi u_n + \Theta\Psi\left(\overline{X}_n - X_n + \tilde{X}_n\right) \\
&= Y_{n+1} - \Theta\Psi\widehat{X}_n - \Theta\Phi u_n.
\end{aligned}$$

Tak więc żądany wzór wynika z pierwszej części. □

Teraz dla funkcjonał u J_N mamy

$$\begin{aligned}
\mathbb{E}\, q_n(X_n, u_n(Y_0,\ldots,Y_n)) &= \mathbb{E}\,\mathbb{E}\left(q_n(X_n, u_n(Y_0,\ldots,Y_n))|Y_0,\ldots,Y_n\right) \\
&= \mathbb{E}\int_{\mathbb{R}^d} q_n(x, u_n(Y_0,\ldots,Y_n))\mu_n(dx),
\end{aligned}$$

gdzie μ_n jest rozkładem gaussowskim z parametrami $\widehat{\overline{X}_n} + \tilde{X}_n$ i P_n. Tak więc funkcjonał J_N możemy zapisać w terminach procesu $(\widehat{X}_n)$. Dokładnie wykazaliśmy następujący fakt.

Propozycja 13.2 *dla dowolnej strategii* π *zachodzi*

$$J_N(\pi, X_0) = \mathbb{E}\left(\sum_{n=0}^{N-1} \widehat{q}_n(\widehat{X}_n, u_n(Y_0,\ldots,Y_n)) + \widehat{r}_N(\widehat{X}_N)\right),$$

gdzie

$$\begin{aligned}
\widehat{q}_n(x,u) &= \int_{\mathbb{R}^d} q_n(y,u)\mathcal{N}(x,P_n)(dy), \quad n = 0,\ldots N-1, \\
\widehat{r}_N(x) &= \int_{\mathbb{R}^d} r_N(y)\mathcal{N}(x,P_N)(dy).
\end{aligned}$$

Z Propozycji 13.1 mamy

$$\widehat{X}_{n+1} = \widehat{F}_n(\widehat{X}_n, u_n, Y_{n+1}),$$

gdzie

$$\widehat{F}_n(x,u,y) = (I - K_{n+1}\Theta)\Psi x + (I - K_{n+1}\Theta)\Phi u + K_{n+1}y.$$

Poniższy rezultat umożliwi nam zastosowanie Twierdzenia 2.3 do rozwiązania problemu sterowania z niepełną informacją.

Propozycja 13.3 *Dla dowolnej strategii π proces $(\widehat{X}, \mathcal{Y}_n)$ jest sterowanym procesem Markowa z rodziną operatorów przejścia*

$$\begin{aligned}\widehat{P}_n\psi(x) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \psi\left(\widehat{F}_n(x,u,\Theta\Psi y + \Theta\Phi u + \Theta\xi + \eta)\right)\\ \times\mathcal{N}(x,P_n)(dy)\mathcal{N}(0,R_1)(d\xi)\mathcal{N}(0,R_2)(d\eta)\end{aligned}$$

Gdzie $\mathcal{N}(0,R_1)$ jest rozkładem ξ_n, $n = 0,1,\ldots$, a $\mathcal{N}(0,R_2)$ jest rozkładem η_n, $n = 0,1,\ldots$.

Dowód Mamy

$$\begin{aligned}\widehat{X}_{n+1} &= \widehat{F}_n(\widehat{X}_n, u_n, Y_{n+1})\\ &= \widehat{F}_n(\widehat{X}_n, u_n, \Theta X_{n+1} + \eta_{n+1})\\ &= \widehat{F}_n(\widehat{X}_n, u_n, \Theta\Psi X_n + \Theta\Phi u_n + \Theta\xi_{n+1} + \eta_{n+1}).\end{aligned}$$

Ponieważ $\{\xi_n\}$ i $\{\eta_n\}$ są niezależne mamy

$$\begin{aligned}&\mathbb{E}\left(\psi(\widehat{X}_{n+1})|\mathcal{Y}_n\right)\\ &= \mathbb{E}\left(\psi(\widehat{F}_n(\widehat{X}_n, u_n, \Theta\Psi X_n + \Theta\Phi u_n + \Theta\xi + \eta_{n+1}))|\mathcal{Y}_n\right)\\ &\qquad \mathcal{N}(0,R_1)(d\xi)\\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \psi\left(\widehat{F}_n(\widehat{X}_n, u_n, \Theta\Psi y + \Theta\Phi u_n + \Theta\xi + \eta)\right)\\ &\qquad \mathcal{N}(\widehat{X}_n, P_n)(dy)\mathcal{N}(0,R_1)(d\xi)\mathcal{N}(0,R_2)(d\eta). \qquad \square\end{aligned}$$

14

Równania filtracji dla skończonego łańcucha Markowa

14.1 Zasadniczy rezultat

Załóżmy, że $X = (X_n)$ jest jednorodnym łańcuchem Markowa na skończonej przestrzeni stanów $E = \{1, \ldots, K\}$. Niech $(p_{i,j})$ będzie macierzą przejścia dla X. Proces X traktujemy jako proces nieobserwowalny. Proces obserwacji (Y_n) zdefiniowany jest następująco

$$Y_n = \xi_n^{X_n}, \qquad n = 0, 1, \ldots,$$

gdzie (ξ_n^k), $k \in E$, $n = 0, 1, \ldots$ są niezależnymi zmiennymi losowymi w zbiorze $G = \{1, \ldots, M\}$. Tak więc zamiast obserwacji $X_n = k$ dostępny jest jedynie sygnał ξ_n^k wysłany ze stanu k. Zakładamy, że rozkład ξ_n^k nie zależy od n. Połóżmy

$$q^k(y) := \mathbb{P}\,(\xi_n^k = y), \qquad k \in E,\ y \in G,\ n = 0, 1, \ldots .$$

Oznaczmy przez Π_n wektor prawdopodobieństwa warunkowego X_n względem Y_n. Dokładnie

$$\Pi_n = \left(\Pi_n^1, \ldots, \Pi_n^K\right),$$

gdzie

$$\Pi_n^k = \mathbb{P}\,(X_n = k | \mathcal{Y}_n), \qquad k \in E,\ n = 0, 1, \ldots,$$
$$\mathcal{Y}_n = \sigma(Y_0, \ldots, Y_n), \qquad n = 0, 1, \ldots .$$

Niech

$$\Delta = \left\{ v = (v_1, \ldots, v_K) \colon v_k \geq 0,\ \sum_k v_k = 1 \right\}$$

oznacza przestrzeń stanów procesu Π. Następujący resultat składa się z dwóch części. Pierwsza to tak zwane *równania filtracji.* Dają one wzór rekurencyjny na Π w terminach obserwacji Y. Z drugiej części wynika Markowskość $(\Pi_n, \mathcal{Y}_n)$.

Twierdzenie 14.1 *Dla* $n = 0, 1, \ldots,$

$$\Pi_{n+1} = F(\Pi_n, Y_{n+1}), \qquad n = 0, 1, \ldots,$$

gdzie $F\colon \Delta \times G \to \Delta,$

$$F(v, y) = (F_1(v, y), \ldots, F(v, y)),$$

zadany jest następująco

$$F_k(v, y) = \frac{q^k(y) \sum_j p_{j,k} v_j}{\sum_{i,j} q^i(y) p_{j,i} v_j}.$$

Rozkład początkowy

$$\Pi_0 = \left(\Pi_0^1, \ldots, \Pi_0^K\right)$$

dany jest wzorem

$$\Pi_0^k = \frac{\mathbb{P}\,(X_0 = k) q^k(Y_0)}{\sum_j \mathbb{P}\,(X_0 = j) q^j(Y_0)}, \qquad j \in E.$$

Ponadto process $(\Pi_n, \mathcal{Y}_n)$, $n = 0, 1, \ldots$ *jest jednorodnym łańcuchem Markowa na* Δ *z operatorem przejścia*

$$P\psi(v) = \sum_{y \in G} \psi(F(v, y)) \sum_{l,k \in E} q^l(y) v_k p_{k,l}.$$

Dowód Skorzystamy z następujących dwóch elementarnych faktów. Mianowicie, jeżeli Z i W są zmiennymi losowymi przyjmującymi skończoną liczbę wartości to, patrz Lemat 1.2,

$$\mathbb{P}\,(Z = z | \sigma(W)) \chi_{\{W=w\}} = \mathbb{P}\,(Z = z | W = w) \chi_{\{W=w\}},$$

oraz

$$\mathbb{P}\,(Z = z | W = w) = \frac{\mathbb{P}\,(\{Z = z\} \cap \{W = w\})}{\mathbb{P}\,(\{W = w\})}.$$

W ostatniej równości przyjmujemy, że $0/0 = 0$.

Przechodzimy do dowodu pierwszej części twierdzenia. Wykorzystując następujące tożsamości

$$\begin{aligned}
\mathbb{P}\,(A | B \cap C) &= \frac{\mathbb{P}\,(A \cap B \cap C)}{\mathbb{P}\,(B \cap C)} = \frac{\mathbb{P}\,(A \cap B \cap C)}{\mathbb{P}\,(B)} \, \frac{\mathbb{P}\,(B)}{\mathbb{P}\,(B \cap C)} \\
&= \frac{\mathbb{P}\,(A \cap C | B)}{\mathbb{P}\,(C | B)}
\end{aligned}$$

otrzymujemy

$$
\begin{aligned}
&\mathbb{P}\left(X_{n+1}=j|Y_0=y_0,\dots,Y_n=y_n,Y_{n+1}=y\right)\\
&=\frac{\mathbb{P}\left(X_{n+1}=j,Y_{n+1}=y|Y_0=y_0,\dots Y_n=y_n\right)}{\mathbb{P}\left(Y_{n+1}=y|Y_0=y_0,\dots,Y_n=y_n\right)}\\
&=\frac{\sum_k \mathbb{P}\left(X_{n+1}=j,X_n=k,Y_{n+1}=y|Y_0=y_0,\dots,Y_n=y_n\right)}{\sum_k \mathbb{P}\left(Y_{n+1}=y,X_n=k|Y_0=y_0,\dots,Y_n=y_n\right)}.
\end{aligned}
$$

Ponieważ ξ^l_{n+1} nie zależy od $Y_0,\dots,Y_n$, X_n i X_{n+1}, mamy

$$
\begin{aligned}
&\mathbb{P}\left(X_{n+1}=l,X_n=m,Y_{n+1}=y|Y_0=y_0,\dots,Y_n=y_n\right)\\
&=\mathbb{P}\left(X_{n+1}=l,X_n=m,\xi^l_{n+1}=y|Y_0=y_0,\dots,Y_n=y_n\right)\\
&=q^l(y)\mathbb{P}\left(X_{n+1}=l,X_n=m|Y_0=y_0,\dots,Y_n=y_n\right).
\end{aligned}
$$

Teraz

$$
\begin{aligned}
&\mathbb{P}\left(X_{n+1}=l,X_n=m|Y_0=y_0,\dots,Y_n=y_n\right)\\
&=\mathbb{P}\left(X_{n+1}=l|Y_0=y_0,\dots,Y_n=y_n,X_n=m\right)\\
&\qquad\times\mathbb{P}\left(X_n=m|Y_0=y_0,\dots,Y_n=y_n\right).
\end{aligned}
$$

Z definicji

$$
\begin{aligned}
&\mathbb{P}\left(X_{n+1}=l|Y_0=y_0,\dots,Y_n=y_n,X_n=m\right)\\
&=\frac{\mathbb{P}\left(X_{n+1}=l,Y_0=y_0,\dots,Y_n=y_n,X_n=m\right)}{\mathbb{P}\left(Y_0=y_0,\dots,Y_n=y_n,X_n=m\right)}.
\end{aligned}
$$

Oczywiście

$$
\begin{aligned}
&\mathbb{P}\left(X_{n+1}=l,Y_0=y_0,\dots,Y_n=y_n,X_n=m\right)\\
&=\sum\mathbb{P}\left(X_{n+1}=l,X_n=m,X_0=x_0,\dots,X_{n-1}=x_{n-1},\right.\\
&\qquad\left.Y_0=y_0,\dots,Y_n=y_n\right),
\end{aligned}
$$

gdzie suma jest po wszystkich ciągach $x_0,\dots,x_n\in E$. Teraz dla ustalonego $x_0,\dots,x_n\in E$ mamy

$$
\begin{aligned}
&\mathbb{P}\left(X_{n+1}=l,X_n=m,X_0=x_0,\dots,X_{n-1}=x_{n-1},\right.\\
&\qquad\left.Y_0=y_0,\dots,Y_n=y_n\right)\\
&=\mathbb{P}\left(X_{n+1}=l,X_n=m,X_0=x_0,\dots,X_{n-1}=x_{n-1},\right.\\
&\qquad\left.\xi^{x_0}_0=y_0,\dots,\xi^{x_{n-1}}_{n-1}=y_{n-1},\xi^m_n=y_n\right).
\end{aligned}
$$

Ponieważ (ξ_j) są niezależne od (X_j) mamy

$$
\begin{aligned}
&\mathbb{P}\left(X_{n+1}=l,X_n=m,X_0=x_0,\dots,X_{n-1}=x_{n-1},\right.\\
&\qquad\left.Y_0=y_0,\dots,Y_n=y_n\right)\\
&=\mathbb{P}\left(X_{n+1}=l,X_n=m,X_0=x_0,\dots,X_{n-1}=x_{n-1}\right)\\
&\qquad\mathbb{P}\left(\xi^{x_0}_0=y_0,\dots,\xi^{x_{n-1}}_{n-1}=y_{n-1},\xi^m_n=y_n\right).
\end{aligned}
$$

Z Markowowskości (X_j) mamy

$$\begin{aligned}&\mathbb{P}\left(X_{n+1}=l, X_n=m, X_0=x_0, \ldots, X_{n-1}=x_{n-1}\right)\\&=\mathbb{P}\left(X_{n+1}=l|X_n=m\right)\mathbb{P}\left(X_n=m, X_0=x_0, \ldots, X_{n-1}=x_{n-1}\right).\end{aligned}$$

Pokazaliśmy więc, że

$$\begin{aligned}&\mathbb{P}\left(X_{n+1}=l, X_n=m|Y_0=y_0, \ldots, Y_n=y_n\right)\\&=\mathbb{P}\left(X_{n+1}=l|Y_0=y_0, \ldots, Y_n=y_n, X_n=m\right)\\&\qquad\times\mathbb{P}\left(X_n=m|Y_0=y_0, \ldots, Y_n=y_n\right)\\&=\mathbb{P}\left(X_{n+1}=l|X_n=m\right)\mathbb{P}\left(X_n=m|Y_0=y_0, \ldots, Y_n=y_n\right)\\&=p_{m,l}\mathbb{P}\left(X_n=m|Y_0=y_0, \ldots, Y_n=y_n\right).\end{aligned}$$

Reasumując

$$\begin{aligned}&\mathbb{P}\left(X_{n+1}=j|Y_0=y_0, \ldots, Y_n=y_n, Y_{n+1}=y\right)\\&=\frac{\sum_k q^j(y)p_{k,j}\mathbb{P}\left(X_n=k|Y_0=y_0, \ldots, Y_n=y_n\right)}{\sum_{k,l} q^l(y)p_{k,l}\mathbb{P}\left(X_n=k|Y_0=y_0, \ldots, Y_n=y_n\right)},\end{aligned}$$

co kończy dowód pierwszej części twierdzenia.

Przechodzimy teraz do dowodu Markowskości procesu $(\Pi_n, \mathcal{Y}_n)$. Ustalmy $\psi \in B_b(\Delta)$. Wówczas, z pierwszej części twierdzenia i z Lematu 1.4 otrzymujemy

$$\begin{aligned}\mathbb{E}\left(\psi(\Pi_{n+1})|\mathcal{Y}_n\right)&=\mathbb{E}\left(\psi(F(\Pi_n, Y_{n+1}))|\mathcal{Y}_n\right)\\&=\sum_{y\in G}\psi(F(\Pi_n, y))\mathbb{P}\left(Y_{n+1}=y|\mathcal{Y}_n\right)\\&=\sum_{y\in G}\psi(F(\Pi_n, y))\sum_{j\in E}q^j(y)\sum_{k\in E}p_{k,j}\Pi_n^k.\end{aligned}$$

□

W rozważaniach dotyczących problemu rozregulowania wykorzystamy następujący techniczny lemat.

Lemat 14.1 *Dla dowolnej funkcji wklęsłej ograniczonej* $\psi\colon \Delta \to [0, +\infty)$ *funkcja* $P\psi$ *jest wklęsła.*

Dowód Z twierdzenia mamy

$$P\psi(v)=\sum_{y\in G}\psi(F(v, y))\sum_{l,k\in E}q^l(y)v_k p_{k,l}.$$

Wystarczy więc wykazać, że dla dowolnej funkcji wklęsłej nieujemnej ψ i dla dowolnego $y \in G$ funkcja

$$\eta\colon \Delta \ni v \to \psi(F(v, y))\sum_{l,k\in E}q^l(y)v_k p_{k,l} \in \mathbb{R}$$

jest wklęsła. Niech

$$r(v) = \sum_j r_j(v), \qquad r_j(v) = \sum_k q^k(y) p_{k,j} v_k,$$

$$R(v) = (r_1(v), \ldots, r_K(v)).$$

Wówczas

$$\eta(v) = \psi\left(\frac{R(v)}{r(v)}\right) r(v).$$

Tak więc dla $v, u \in \Delta$ i dla $\alpha, \beta \geq 0$, $\alpha + \beta = 1$ mamy

$$\begin{aligned}
&\psi\left(\frac{R(\alpha v + \beta u)}{r(\alpha v + \beta u)}\right) r(\alpha v + \beta u) \\
&= \psi\left(\frac{R(v)}{r(v)} \frac{r(\alpha v)}{r(\alpha v) + r(\beta u)} + \frac{R(u)}{r(u)} \frac{r(\beta u)}{r(\alpha v) + r(\beta u)}\right) r(\alpha v + \beta u) \\
&= \alpha\psi\left(\frac{R(v)}{r(v)}\right) r(v) + \beta\psi\left(\frac{R(u)}{r(u)}\right) r(u) \\
&\qquad + \beta\psi\left(\frac{R(v)}{r(v)}\right) r(u) + \alpha\psi\left(\frac{R(u)}{r(u)}\right) r(v) \\
&\geq \alpha\psi\left(\frac{R(v)}{r(v)}\right) r(v) + \beta\psi\left(\frac{R(u)}{r(u)}\right) r(u). \qquad \square
\end{aligned}$$

14.2 Rozwiązanie problemu rozregulowania

Przypomnijmy, że jakość produktu maszyny waha się od 1 do M. Jeżeli maszyna jest w dobrej kondycji to prawdopodbieństwo produkcji wyrobu o jakości $k \in G = \{1, 2, \ldots M\}$ wynosi $q^0(k)$. Gdy maszyna jest w złym stanie to prawdopodobieństwo tego zdarzenia wynosi $q^1(k)$. Po losowym czasie σ maszyna rozregulowuje się. Celem naszym jest znalezienie optymalnego momentu zatrzymania maszyny $\widehat{\tau}$ w celu jej naprawienia. Moment $\widehat{\tau}$ powinien minimalizować funkcjonał kosztów

$$J(\tau) = \mathbb{P}(\tau < \sigma) + c\,\mathbb{E}\left\{(\tau - \sigma) \vee 0\right\}.$$

Przypomnijmy, że część pierwsza funkcjonału reprezentuje kare za opóźnienie, a druga to kara za przedwczesne wstrzymanie produkcji.

Naszym pierwszym celem będzie sprowadzenie problemu do standartowego zagadnienie optymalnego stopowania łańcuchem Markowa. W tym celu niech $X_n \in \{0, 1\}$ oznacza stan maszyny w chwili n. To znaczy $X_n = 0$ maszyna jest w dobrym stanie, a $X_n = 1$ gdy maszyna wymaga naprawy. Zakładamy, że (X_n) jest łańcuchem Markowa ze stanem pochłaniającym $\{1\}$. Niech $(p_{i,j})$ będzie macierzą przejście dla (X_n). To znaczy

$$p_{0,0} = p, \quad p_{0,1} = 1 - p, \quad p_{1,0} = 0, \quad p_{1,1} = 1,$$

gdzie $p \in [0,1]$ jest zadanym parametrem. Niech

$$\left(\xi_0^0, \xi_1^0, \ldots\right) \qquad \text{i} \qquad \left(\xi_0^1, \xi_1^1, \ldots\right)$$

będą ciągami niezależnych zmiennych losowych w G o rozkładach odpowiednio q^0 i q^1. Miech

$$Y_n = \xi_n^{X_n}, \qquad \mathcal{Y}_n = \sigma(Y_0, Y_1, \ldots, Y_n), \qquad n = 0, 1, \ldots.$$

Wówczas τ powinien być momentem Markowa względem filtracji $(\mathcal{Y}_n)$. Niech Σ będzie klasą wszystkich momentów Markowa τ względem $(\mathcal{Y}_n)$ spełniających

$$\mathbb{P}\left(\tau < \infty\right) = 1.$$

Oczywiście

$$\sigma = \inf\{n \colon X_n = 1\} \notin \Sigma.$$

Niech

$$\Pi_n = \mathbb{P}\left(X_n = 1 | \mathcal{Y}_n\right), \qquad n = 0, 1, \ldots.$$

(Π_n) jest procesem na przestrzeni stanów $E = [0,1]$. Niech $B_b([0,1])$ oznacza klasę wszystkich funkcji mierzalnych i ograniczonych na $[0,1]$.

Z Twierdzenia 14.1, $(\Pi_n, \mathcal{Y}_n)$ jest jednorodnym procesem Markowa. Następny fakt zapewnia pożądaną postać funkcjonału kosztu.

Lemat 14.2 *Funkcjonał kosztu dany jest formułą*

$$J(\tau) = \mathbb{E}\left(\sum_{n=0}^{\tau-1} c\,\Pi_n + (1 - \Pi_\tau)\right), \qquad \tau \in \Sigma.$$

Dowód Mamy

$$\mathbb{P}\left(\tau < \sigma\right) = \mathbb{P}\left(X_\tau = 0\right) = \sum_{n=0}^{+\infty} \mathbb{P}\left(X_n = 0, \tau = n\right).$$

Ponieważ $\{\tau = n\} \in \mathcal{Y}_n$, mamy

$$\begin{aligned}
\mathbb{P}\left(\tau < \sigma\right) &= \sum_{n=0}^{+\infty} \mathbb{E}\left(\mathbb{P}\left(X_n = 0 | \mathcal{Y}_m\right); \tau = n\right) \\
&= \sum_{n=0}^{+\infty} \mathbb{E}\left((1 - \Pi_n)\chi_{\{\tau = n\}}\right) \\
&= \mathbb{E}\left(1 - \Pi_\tau\right).
\end{aligned}$$

Zauważmy również, że

$$(\tau - \sigma) \vee 0 = \sum_{n=0}^{\tau-1} X_n = \sum_{n=0}^{+\infty} X_n \chi_{\{\tau > n\}}.$$

Stąd

$$\begin{aligned}\mathbb{E}\,(\tau-\sigma)\vee 0 &= \sum_{n=0}^{+\infty}\mathbb{E}\,X_n\chi_{\{\tau>n\}}\\ &= \sum_{n=0}^{+\infty}\mathbb{E}\,\mathbb{E}\left(X_n\chi_{\{\tau>n\}}|\mathcal{Y}_n\right)\end{aligned}$$

Ponieważ $\tau > n$ jest $\mathcal{Y}_n$-mierzalne, mamy

$$\begin{aligned}\mathbb{E}\,(\tau-\sigma)\vee 0 &= \sum_{n=0}^{+\infty}\mathbb{E}\left(\mathbb{E}\,(X_n|\mathcal{Y}_n)\,\chi_{\{\tau>n\}}\right)\\ &= \sum_{n=0}^{+\infty}\mathbb{E}\left(\mathbb{E}\,(\chi_{X_n}|\mathcal{Y}_n)\,\chi_{\{\tau>n\}}\right)\\ &= \sum_{n=0}^{+\infty}\mathbb{E}\left(\mathbb{P}\,(X_n=1|\mathcal{Y}_n)\,\chi_{\{\tau>n\}}\right)\\ &= \mathbb{E}\sum_{n=0}^{\tau-1}\Pi_n. \qquad \square\end{aligned}$$

Reasumując sprowadziliśmy problem rozregulowania do problemu stopowania na nieskończonym przedziale czasowym procesu Markowa $(\Pi_n, \mathcal{Y}_n)$ z funkcjonałem kosztu

$$J(\tau) = \mathbb{E}\left(\sum_{n=0}^{\tau-1} q(\Pi_n) + r(\Pi_\tau)\right),$$

gdzie

$$q(x) = cx, \qquad r(x) = 1-x, \qquad\qquad x \in E = [0,1].$$

Aby rozwiązać problem rozregulowania skorzystamy z Twierdzenia 5.4. Niech P będzie operatorem przejścia dla $(\Pi_n, \mathcal{Y}_n)$. Dla $x \in E = [0,1]$ niech

$$\begin{aligned}v_0(x) &= r(x) = 1-x,\\ v_{n+1}(x) &= \min\left\{q(x) + Pv_n(x), r(x)\right\}\\ &= \min\left\{cx + Pv_n(x), 1-x\right\}, \quad n = 0,\ldots.\end{aligned}$$

Ponieważ, patrz Lemat 14.1, P przeprowadza zbiór funkcji wklęsłych nieujemnych w siebie wszystkie funkcje v_n, $n = 0,\ldots$, są wklęsłe nieujemne. Funkcje wklęsłe na przedziale ograniczonym mogą być nieciągłe tylko na brzegach. Ponieważ $r(1) = 0$, mamy $v_n(1) = 0$. Ponieważ

$$0 \le v_n(x) \le 1-x = r(x), \qquad x \in [0,1],$$

funkcje (v_n) są ciągłe w 1. Z wklęsłości i ciągłości wynika, że istnieje ciąg (a_n) liczb z przedziału $[0,1]$, taki że

$$\{x \in [0,1]\colon v_n(x) \geq r(x)\} = [a_n, 1], \qquad n = 0, 1, \ldots.$$

Ponieważ (v_n) jest ciągiem malejącym, ciąg (a_n) rośnie. Niech

$$a = \lim_{n\to+\infty} a_n.$$

Pokażemy, że

$$\widehat{\tau} = \min\{n\colon \Pi_n \geq a\}$$

jest optymalnym momentem zatrzymania. W tym celu oznaczmy przez Σ_N, $N = 1, 2, \ldots$ klasę momentów Markowa, względem $(\mathcal{Y}_n)$ spełniających

$$\mathbb{P}(\tau \leq N) = 1.$$

Z Twierdzenia 5.4,

$$\inf_{\tau\in\Sigma_N} J(\tau) = \mathbb{E}\, v_N(\Pi_0) = J(\widehat{\tau}_N),$$

gdzie

$$\widehat{\tau}_N = \min\{n \leq N\colon \Pi_n \geq a_{N-n}\}, \qquad N = 1, \ldots.$$

Stąd, ponieważ (v_n) są wspólnie ograniczone przez 1,

$$\begin{aligned}
\inf_{\tau\in\Sigma} J(\tau) &= \inf_N \inf_{\tau\in\Sigma_N} J(\tau) \\
&= \inf_N \mathbb{E}\, v_N(\Pi_0) = \lim_{n\to+\infty} \mathbb{E}\, v_N(\Pi_0) \\
&= \lim_{N\to+\infty} J(\widehat{\tau}_N) = \mathbb{E}\, v_\infty(\Pi_0),
\end{aligned}$$

gdzie v_∞ jest granicą punktową ciągu malejącego $\{v_N\}$. Wystarczy więc pokazać, że

$$\lim_{N\to+\infty} J(\widehat{\tau}_N) = J(\widehat{\tau}).$$

Biorąc pod uwage nieujemność q i r wystarczy zauważyć, że

$$\widehat{\tau}_N \uparrow \widehat{\tau}.$$

Mamy więc wzór na optymalny moment zatrzymania w terminach (Π_n). Z Twierdzenia 14.1 wyliczamy (Π_n) w terminach procesu obserwacji (Y_n).

Część II

Czas ciągły

15

Sterowanie w czasie ciągłym

15.1 Układ deterministyczny

Zakładamy, że zadane są otwarty podzbiór $E \subseteq \mathbb{R}^d$, przestrzeń mierzalna $(U,\mathcal{U})$, odwzorowanie mierzalne $f\colon E \times U \to \mathbb{R}^d$ oraz funkcje mierzalne $g\colon \times U \to \mathbb{R}$ i $G\colon E \to \mathbb{R}$.

Ustalmy $T \in (0, +\infty)$. Rozważmy układ sterowany

$$\frac{dy}{dt} = f(y, u), \qquad y(0) = x, \tag{15.1}$$

z funkcjonałem kosztu

$$J_T(x, u) = \int_0^T g(y(s), u(s))ds + G(y(T)).$$

Odwzorowanie mierzalne $u\colon [0, T] \to U$ jest *sterowaniem dopuszczalnym* gdy dla dowolnego x równanie (15.1) ma dokładnie jedno rozwiązanie $y^{x,u}(t)$, $t \in [0, T]$.

Celem jest znalezienie optymalnego sterowania przy ustalonym warunku początkowym x, czyli sterowania dopuszczalnego $\widehat{u}$, takiego że dla dowolnego sterowania dopuszczalnego u zachodzi

$$J_T(x, \widehat{u}) \leq J_T(x, u).$$

Twierdzenie 15.1 *Niech V będzie funkcją ciągłą na $[0, T] \times E$, klasy C^1 na $(0, T) \times E$ spełniającą na $(0, T) \times E$ równanie*

$$\frac{\partial V}{\partial t}(t, x) = \inf_{u \in U} \left(g(x, u) + \langle \nabla_x V(t, x), f(x, u)\rangle\right) \tag{15.2}$$

z warunkiem początkowym

$$V(0, x) = G(x), \qquad x \in E.$$

Wówczas dla dowolnego $x \in E$ i dowolnego sterowania dopuszczalnego u zachodzi

$$J_T(x,u) \geq V(T,x). \tag{15.3}$$

Przypuśćmy, że dla pewnego mierzalnego odwzorowania

$$\widehat{v}\colon [0,T] \times E \to U$$

mamy

$$\begin{aligned} &g(x,\widehat{v}(t,x)) + \langle \nabla_x V(t,x), f(t,\widehat{v}(t,x)\rangle \\ &\leq g(x,u) + \langle \nabla_x V(t,x), f(t,u) \rangle, \qquad t \in (0,T),\ x \in E,\ u \in U. \end{aligned}$$

Ponadto, załóżmy, że równanie

$$\begin{aligned} \frac{dy}{dt}(t) &= f(y(t),\widehat{v}(T-t,y(t)), \quad t \in (0,T), \\ y(0) &= x \end{aligned}$$

ma dokładnie jedno rozwiązanie $\widehat{y}$, ciągłe w $[0,T]$ i absolutnie ciągłe w $(0,T)$, o wartościach w E. Wówczas sterowanie

$$\widehat{u}(t) = \widehat{v}(T-t,\widehat{y}(t)), \qquad t \in [0,T]$$

jest optymalne i

$$J_T(x,\widehat{u}) = V(T,x).$$

Dowód Niech u będzie sterowaniem dopuszczalnym oraz niech y będzie odpowiadającym mu rozwiązaniem. Niech

$$\omega(t) = V(T-t,y(t)), \qquad t \in [0,T].$$

Jest to funkcja ciągła, absolutnie ciągła na dowolnym przedziale $[a,b]$ zawartym w $(0,T)$. Ponadto dla prawie każdego $t \in (0,T)$ mamy

$$\begin{aligned} \frac{d\omega}{dt}(t) &= -\frac{\partial V}{\partial t}(T-t,y(t)) + \left\langle \nabla_x V(T-t,y(t)), \frac{dy}{dt}(t) \right\rangle \\ &= -\frac{\partial V}{\partial t}(T-t,y(t)(+ \langle \nabla_x V(T-t,y(t)), f(y(t),u(t)) \rangle . \end{aligned}$$

Stąd mamy

$$\begin{aligned} &V(T-b,y(b)) - V(T-a,y(a)) = \omega(b) - \omega(a) \\ &= \int_a^b \frac{d\omega}{\mathrm{dt}}(t)dt \\ &= \int_a^b \left\{ -\frac{\partial V}{\partial t}(T-t,y(t)) + \langle \nabla_x V(T-t,y(t)), f(y(t),u(t))) \rangle \right\} dt \\ &\geq -\int_a^b g(y(t),u(t))dt. \end{aligned}$$

Przechodząc do granicy $a \to 0$, $b \to T$ otrzymujemy

$$G(y(T)) - V(T,x) \geq -\int_0^T g(y(t),u(t))dt$$

czyli żądaną nierówność (15.3). W ten sam sposób dochodzimy do równości

$$\begin{aligned}
&G(\widehat{y}(T)) - V(T,x) \\
&= \int_0^T \left\{ -\frac{\partial V}{\partial t}(T-t,\widehat{y}(t)) + \langle \nabla_x V(T-t,\widehat{y}(t)), f(t,\widehat{y}(t))\rangle \right\} dt \\
&= -\int_0^T g(\widehat{y}(t),\widehat{u}(t))dt.
\end{aligned}$$

Czyli

$$J_T(x,\widehat{u}) = G(\widehat{y}(T)) + \int_0^T g(\widehat{y}(t),\widehat{u}(t))dt = V(T,x). \qquad \Box$$

Z dowolnym równaniem

$$\frac{dy}{dt} = f(y), \qquad y(0) = x \in E \tag{15.4}$$

można związać następujący operator różniczkowy

$$L\psi(x) = \langle \nabla_x \psi(x), f(x)\rangle, \qquad x \in E,$$

który nazywać będziemy *generatorem układu* (15.4).

Twierdzenie 15.2 *Jeżeli $y^x(t)$, $t \geq 0$ jest rozwiązaniem* (15.4) *to przy założeniu, że G i g są funkcjami klasy C^1, funkcja*

$$V(t,x) = G(y^x(t)) + \int_0^t g(y^x(s))ds, \qquad t \geq 0$$

jest rozwiązaniem równania

$$\frac{\partial V}{\partial t}(t,x) = g(x) + LV(t,x), \qquad t > 0,\ x \in E \tag{15.5}$$

z warunkiem początkowym $V(0,x) = G(x)$, $x \in E$.

Dowód Niech

$$V_1(t,x) = g(y^x(t)), \qquad V_2(t,x) = G(y^x(t)), \qquad t \geq 0,\ x \in E.$$

Wtedy

$$\begin{aligned}\frac{\partial V_1}{\partial t}(t,x) &= \left\langle \nabla g(y^x(t)), \frac{d}{dt}y^x(t) \right\rangle \\ &= \langle \nabla g(y^x(t)), f(y^x(t)) \rangle \\ &= Lg(y^x(t)), \qquad t \geq 0.\end{aligned}$$

W szczególności

$$\frac{\partial V}{\partial t}(0,x) = Lg(x) = LV_1(0,x), \qquad x \in E.$$

Ponieważ

$$\begin{aligned}V_1(s+t,x) = g(y^x(s+t)) &= g\left(y^{y^x(s)}(t)\right) \\ &= V_1(t, y^x(s)),\end{aligned}$$

zachodzi

$$\frac{\partial^+}{\partial t} V_1(t,x) = LV_1(t,x), \qquad t \geq 0,\ x \in E.$$

Stąd i z ciągłości funkcji LV_1 wynika, że istnieje również pochodna lewostronna oraz

$$\frac{\partial^-}{\partial t} V_1(t,x) = LV_1(t,x), \qquad t \geq 0,\ x \in E.$$

Stąd istnieje pochodna dwustronna oraz

$$\frac{\partial V_1}{dt}(t,x) = LV_1(t,x), \qquad t \geq 0,\ x \in E.$$

Podobnie otrzymujemy

$$\frac{\partial V_2}{dt}(t,x) = LV_2(t,x), \qquad t \geq 0,\ x \in E.$$

Zauważmy, że

$$V(t,x) = V_2(t,x) + \int_0^t V_1(s,x)ds.$$

Stąd

$$\begin{aligned}LV(t,x) &= LV_2(t,x) + \int_0^t LV_1(s,x)ds \\ &= \frac{\partial V_2}{\partial t}(t,x) + \int_0^t \frac{\partial V_1}{\partial s}(s,x)ds \\ &= \frac{\partial V_2}{\partial t}(t,x) + V_1(t,x) - g(x).\end{aligned}$$

Czyli

$$\frac{\partial V}{\partial t}(t,x) = \frac{\partial V_2}{\partial t}(t,x) + V_1(t,x).$$

Zachodzi więc (15.5). □

Niech teraz L^u będzie operatorem związanym z równaniem

$$\frac{dy}{dt} = f(y,u), \qquad y(0) = x,$$

w którym prawa strona zależy od $u \in U$. Równanie Bellmana możemy zapisać w postaci

$$\frac{\partial V}{\partial t}(t,x) = \inf_{u\in U}\left(g(x,u) + L^u V(t,x)\right), \qquad t > 0,\ x \in E, \tag{15.6}$$

z warunkiem początkowym $V(0,x) = G(x)$, $x \in E$. Zauważmy, że gdy U jest jednoelementowy to (15.6) jest identyczny z równaniem (15.5).

15.2 Proces Wienera

Procesem Wienera z macierzą kowariancji $Q = (q_{i,j})$ nazywamy proces gaussowski $W(t) = (W_1(t), \ldots, W_d(t))$ spełniający

$$\mathbb{E}\, W_j(t) = 0, \qquad \mathbb{E}\, W_j(t) W_i(s) = t \wedge s\, q_{j,i}.$$

Dodatkowo zakłada się, że $W(0) = 0$ i ciągłość trajektorii procesu W. Zauważmy, że dla $s \geq t \geq u \geq r \geq 0$ oraz $i,j = 1, \ldots d$ mamy

$$\mathbb{E}\left(W_j(s) - W_j(t)\right)\left(W_i(u) - W_i(r)\right) = 0.$$

Dlatego przyrosty procesu są nieskorelowane. Ponieważ proces jest gaussowski, z nieskorelowania wynika niezależność przyrostów.

Trajektorie procesu Wienera są funkcjami nieróżniczkowalnymi w sensie klasycznym. Załóżmy, że $d = 1$. Wówczas dla dowolnego $T \in (0, +\infty)$ oraz dla dowolnych funkcji testujących $\psi, \phi \in C_0^\infty([0,T])$ mamy

$$\begin{aligned}
&\mathbb{E}\left(\frac{d}{dt}W, \psi\right)\left(\frac{d}{dt}W, \phi\right) \\
&=: \mathbb{E}\int_0^T\int_0^T W(s)\psi'(s)W(t)\phi'(t)dsdt \\
&= \int_0^T\int_0^T t\wedge s\psi'(t)\phi'(s)dtds \\
&= \int_0^T \psi'(t)\left(\int_0^t s\phi'(s)ds + t\int_t^T \phi'(s)ds\right)dt \\
&= \int_0^T \psi'(t)\int_0^t s\phi'(s)dsdt + \int_0^T \psi'(t)t\int_t^T \phi'(s)dsdt \\
&= \int_0^T \psi(t)\phi(t)dt.
\end{aligned}$$

W szczególności gdy ψ i ϕ mają rozłączne nośniki to

$$\left(\frac{d}{dt}W, \psi\right) \quad \text{oraz} \quad \left(\frac{d}{dt}W, \phi\right)$$

są niezależnymi zmiennymi losowymi.

Formalnie, równanie

$$\frac{dy}{dt} = F(y) + \frac{dW}{dt}$$

opisuje sytuacje, w której szum ma dużą zmienność w czasie, a na rozłącznych przedziałach czasowych zaburzenia są niezależnymi zmiennymi losowymi.

15.3 Układ stochastyczny

Rozważmy stochastyczny układ sterowany

$$\frac{dy}{dt} = f(y(t), u(t)) + B\frac{dW}{dt}(t), \qquad y(0) = x, \tag{15.7}$$

gdzie W jest r-wymiarowym procesem Wienera, a B macierzą wymiaru $d \times r$. Równanie (15.7) rozumiemy jako równanie całkowe

$$y(t) = x + \int_0^t f(y(s), u(s))ds + BW(t), \qquad t \geq 0.$$

Niech $(\mathcal{F}_t)$ będzie wstępującą rodziną σ-podciał, taką że $W(t)$ jest $\mathcal{F}_t$-mierzalny oraz $W(t) - W(s)$, $t \geq s \geq 0$ są niezależne od $\mathcal{F}_s$. Przez sterowanie rozumiemy dowolny adoptowalny proces $u(t)$, $t \geq 0$, którego trajektorie są lokalnie ograniczone. Zakładamy, że przestrzeń parametrów sterujących U jest równa $\mathbb{R}^l$. Aby móc swobodnie korzystać z twierdzenia Fubiniego zakładamy dodatkowo, że dla dowolnego $T > 0$ funkcja

$$u(t, \omega), \qquad t \in [0, T],\ \omega \in \Omega,$$

jest mierzalna ze względu na σ-ciał o $\mathcal{B}([0,T]) \times \mathcal{F}_T$. Czyli zakładamy tak zwaną *progresywną mierzalność u*.

Twierdzenie 15.3 *Załóżmy, że $E = \mathbb{R}^d$ oraz, że dla dowolnego podzbioru ograniczonego $\tilde{U} \subset U$ istnieje stała $K > 0$, taka że*

$$|f(x,u) - f(y,u)| \leq K|x - y|, \qquad x, y \in E,\ u \in \tilde{U}.$$

Wtedy dla dowolnej funkcji ciągłej ψ o wartościach w E i dla dowolnej mierzalnej lokalnie ograniczonej funkcji u istnieje dokładnie jedna funkcja ciągła $v \in C([0,\infty); \mathbb{R}^d)$, taka że

$$v(t) = \int_0^t f(v(s), u(s))ds + \psi(t), \qquad t \geq 0.$$

Dowód Ustalmy $T > 0$ i funkcje u. Rozpatrzmy odwzorowanie

$$\mathcal{G}\colon C([0,T];\mathbb{R}^d) \to C([0,T];\mathbb{R}^d)$$

dane wzorem

$$\mathcal{G}(v)(t) = \int_0^t f(v(s), u(s))ds + \psi(t), \qquad t \in [0,T].$$

Dla dowolnego $\lambda \geq 0$ zdefiniujemy norme Bieleckiego

$$||v||_\lambda = \sup_{0\leq t\leq T} e^{-\lambda t}|v(t)|$$

równoważną tradycyjnej normie $||\cdot||_0$. Mamy

$$\begin{aligned}|\mathcal{G}(v)(t) - \mathcal{G}(z)(t)| &\leq K\int_0^t |v(s) - z(s)|ds \\ &\leq K\int_0^t e^{\lambda s}e^{-\lambda s}|v(s) - z(s)|ds \\ &\leq K\int_0^t e^{\lambda s}ds||v - z||_\lambda \\ &\leq K\frac{e^{\lambda t}}{\lambda}||v - z||_\lambda.\end{aligned}$$

Stąd

$$||\mathcal{G}(v) - \mathcal{G}(z)||_\lambda \leq \frac{K}{\lambda}||v - z||_\lambda.$$

Czyli dla $\lambda > K$ odwzorowanie $\mathcal{G}$ jest zwężające, a więc równanie

$$v = \mathcal{G}(v) + \psi$$

ma dokładnie jedno rozwiązanie. □

Rozpatrzmy w szczególności równanie

$$y(t) = x + \int_0^t f(y(s))ds + BW(t), \qquad t \geq 0, \tag{15.8}$$

gdzie funkcja f nie zależy od parametru sterującego. Rozwiązanie (15.8) będziemy oznaczać przez $y^x(t)$, $t \geq 0$, $x \in \mathbb{R}^d$. Jeżeli ξ jest zmienną losową o wartościach w $\mathbb{R}^d$ mierzalną względem $\mathcal{F}_0$, to $y^\xi(t)$, $t \geq 0$ oznacza rozwiązanie równania, w którym x został o zastąpione przez ξ.

Twierdzenie 15.4 *Załóżmy, że funkcja g wraz z pierwszą i drugą pochodną jest ciągła i ograniczona, to jest, że $g \in C_b^2$. Wtedy funkcja*

$$v(t,x) = \mathbb{E}\, g(y^x(t)), \qquad t \geq 0,\ x \in \mathbb{R}^d$$

jest rozwiązaniem równania

$$\begin{aligned}\frac{\partial v}{\partial t}(t,x) &= Lv(t,x), \qquad t>0,\ x\in\mathbb{R}^d,\\ v(0,x) &= g(x),\end{aligned}$$

gdzie

$$L\psi(x) = \frac{1}{2}\mathrm{Trace}\, QD^2\psi(x) + \langle \nabla\psi(x), f(x)\rangle, \qquad x\in\mathbb{R}^d$$

oraz $Q = BB^*$.

Rozpatrzmy teraz problem znajdywania optymalnego sterowania dla funkcjonału kosztu

$$J_T(x,u) = \mathbb{E}\left(\int_0^T g(y^{x,u}(s),u(s))ds + G(y^{x,u}(T))\right).$$

Oznaczmy przez L^u operator

$$L\psi(x) = \frac{1}{2}\mathrm{Trace}\, QD^2\psi(x) + \langle \nabla\psi(x), f(x,u)\rangle, \qquad x\in\mathbb{R}^d$$

Równanie Bellmana ma wtedy postać

$$\begin{aligned}\frac{\partial V}{\partial t}(t,x) &= \inf_{u\in U}\left(q(x,u) + L^u V(t,x)\right), \qquad t>0,\ x\in\mathbb{R}^d,\\ V(0,x) &= G(x).\end{aligned}$$

16

Dowód twierdzenia o stopowaniu

Niech $(E, \mathcal{E})$ będzie przestrzenią mierzalną. Oznaczmy przez $B_b(E)$ ogół funkcji $\psi\colon E \to \mathbb{R}$ mierzalnych i ograniczonych. Niech $(X_n, \mathcal{F}_n), n = 0, 1, \ldots$ będzie jednorodnym procesem Markowa o wartościach w E, względem filtracji $(\mathcal{F}_n)$ z operatorem przejścia P. Przypomnijmy, że oznacza to w szczególności, że

$$\mathbb{E}\left(\psi(X_{n+1})|\mathcal{F}_n\right) = P\psi(X_n)$$

dla dowolnych $n \in \mathbb{N}$ i $\psi \in B_b(E)$.

Niech Σ_N będzie klasą wszystkich momentów Markowa

$$\tau\colon \Omega \to \{0, 1, \ldots, N\}$$

względem filtracji $(\mathcal{F}_n)$.

Dla zadanych: skończonego horyzontu czasowego $N \in \mathbb{N}$, mierzalnych funkcji $r, q\colon E \to [0, +\infty)$ oraz warunku początkowego X_0, maksymalizujemy (odp. minimalizujemy) średni zysk (odp. koszt)

$$J_{N,r}(\tau, X_0) = \mathbb{E}\left(\sum_{n=0}^{\tau-1} q(X_n) + r(X_\tau)\right),$$

po $\tau \in \Sigma_N$.

Dla problemu z zyskiem definiujemy

$$Q\psi(x) = \max\{q(x) + P\psi(x), r(x)\}, \qquad \psi \in B_b(E),\ x \in E.$$

Dla kosztu zamiast maksimum rozważamy minimum. W następującym twierdzeniu rozważamy funkcjonał zysku.

Twierdzenie 16.1 (i) *Dla dowolnego* $\tau \in \Sigma_N$ *zachodzi*

$$J_{N,r}(\tau, X_0) \le \mathbb{E}\, Q^N r(X_0).$$

(ii) *Moment Markowa*

$$\hat{\tau} = \min\{n \colon Q^{N-n} r(X_n) = r(X_n)\} \tag{16.1}$$

jest optymalny. Ponadto

$$J_{N,r}(\hat{\tau}, X_0) = \mathbb{E}\, Q^N r(X_0). \tag{16.2}$$

Dowód Ustalmy $\tau \in \Sigma_N$. Wówczas

$$\begin{aligned}
&J_{N,r}(\tau, X_0) = \mathbb{E}\left(\sum_{n=0}^{\tau-1} q(X_n) + r(X_\tau)\right)\\
&= \mathbb{E}\,\chi_{\{\tau=N\}}\left(\sum_{n=0}^{\tau-1} q(X_n) + r(X_\tau)\right) + \mathbb{E}\,\chi_{\{\tau<N\}}\left(\sum_{n=0}^{\tau-1} q(X_n) + r(X_\tau)\right)\\
&= \mathbb{E}\,\chi_{\{\tau>N-1\}}\left(\sum_{n=0}^{N-1} q(X_n) + r(X_N)\right) + \mathbb{E}\,\chi_{\{\tau<N\}}\left(\sum_{n=0}^{\tau-1} q(X_n) + r(X_\tau)\right).
\end{aligned}$$

Następnie

$$\begin{aligned}
&\mathbb{E}\,\chi_{\{\tau>N-1\}}\left(\sum_{n=0}^{N-1} q(X_n) + r(X_N)\right)\\
&= \mathbb{E}\,\mathbb{E}\left(\chi_{\{\tau>N-1\}}\left(\sum_{n=0}^{N-1} q(X_n) + r(X_N)\right)\Big|\mathcal{F}_{N-1}\right)\\
&= \mathbb{E}\,\chi_{\{\tau>N-1\}}\left(\sum_{n=0}^{N-1} q(X_n) + \mathbb{E}\left(r(X_N)|\mathcal{F}_{N-1}\right)\right)\\
&= \mathbb{E}\,\chi_{\{\tau>N-1\}}\left(\sum_{n=0}^{N-1} q(X_n) + Pr(X_{N-1})\right)\\
&= \mathbb{E}\,\chi_{\{\tau>N-1\}}\left(\sum_{n=0}^{N-2} q(X_n) + q(X_{N-1}) + Pr(X_{N-1})\right)\\
&= \mathbb{E}\,\chi_{\{\tau>N-1\}}\sum_{n=0}^{\tau\wedge(N-1)-1} q(X_n)\\
&\qquad + \mathbb{E}\,\chi_{\{\tau>N-1\}}\left(q(X_{N-1}) + Pr(X_{N-1})\right)\\
&= \mathbb{E}\,\chi_{\{\tau>N-1\}}\sum_{n=0}^{\tau\wedge(N-1)-1} q(X_n)\\
&\qquad + \mathbb{E}\,\chi_{\{\tau>N-1\}}\left(q(X_{\tau\wedge(N-1)}) + Pr(X_{\tau\wedge(N-1)})\right).
\end{aligned}$$

Ponadto

$$
\begin{aligned}
&\mathbb{E}\,\chi_{\{\tau<N\}}\left(\sum_{n=0}^{\tau-1} q(X_n)+r(X_\tau)\right)\\
&\quad=\mathbb{E}\,\chi_{\{\tau<N\}}\left(\sum_{n=0}^{\tau\wedge(N-1)-1} q(X_n)+r(X_{\tau\wedge(N-1)})\right).
\end{aligned}
$$

Reasumując

$$
\begin{aligned}
J_{N,r}(\tau,X_0)&=\mathbb{E}\sum_{n=0}^{\tau\wedge(N-1)-1} q(X_n)\\
&+\mathbb{E}\,\chi_{\{\tau>N-1\}}\big(q(X_{\tau\wedge(N-1)})+Pr(X_{\tau\wedge(N-1)})\big)\\
&+\mathbb{E}\,\chi_{\{\tau<N\}}r(X_{\tau\wedge(N-1)}).
\end{aligned}
$$

Ponieważ dla funkcjonału zysku

$$
Qr(X_{\tau\wedge(N-1)})\geq r(X_{\tau\wedge(N-1)})
$$

oraz

$$
Qr(X_{\tau\wedge(N-1)})\geq q(X_{\tau\wedge(N-1)})+Pr(X_{\tau\wedge(N-1)}),
$$

dla kosztu mamy nierówności przeciwne, otrzymujemy

$$
\begin{aligned}
&J_{N,r}(\tau,X_0)\\
&\leq\mathbb{E}\left(\sum_{n=0}^{\tau\wedge(N-1)-1} q(X_n)+Qr(X_{\tau\wedge(N-1)})\right)=J_{N-1,Qr}(\tau\wedge(N-1),X_0).
\end{aligned}
$$

Stąd iterując N razy otrzymujemy

$$
J_{N,r}(\tau,X_0)\leq J_{N-N,Q^N r}(\tau\wedge(N-N),X_0)=\mathbb{E}\,Q^N r(X_0),
$$

co dowodzi pierwszą część twierdzenia.

Dowód drugiej części będzie zakończony jak tylko wykażemy, że dla $\hat{\tau}$ danego przez (16.1) zachodzi (16.2). Dowód będzie indukcyjny, ze względu na horyzont czasowy N. Dla $N=0$ teza jest oczywista. Załóżmy, że (16.2) zachodzi dla horyzontu $N-1$. Rozważmy deterministyczny X_0 spełniający

$$
Q^N r(X_0)>r(X_0).
$$

Wówczas $\mathbb{P}(\hat{\tau}=0)=0$. Mamy

$$
J_{N,r}(\hat{\tau},X_0)=q(X_0)+\mathbb{E}\left(\sum_{n=1}^{\hat{\tau}-1} q(X_n)+r(X_{\hat{\tau}})\right).
$$

Oczywiście

$$\mathbb{E}\left(\sum_{n=1}^{\hat\tau-1} q(X_n) + r(X_{\hat\tau})\right) = J_{N-1,r}(\hat\tau, X_1),$$

a więc z założenia indukcyjnego

$$\mathbb{E}\left(\sum_{n=1}^{\hat\tau-1} q(X_n) + r(X_{\hat\tau})\right) = \mathbb{E}\, Q^{N-1} r(X_1).$$

Tak więc

$$\begin{aligned} J_{N,r}(\hat\tau, X_0) &= q(x_0) + \mathbb{E}\, Q^{N-1} r(X_1) \\ &= q(x_0) + \mathbb{E}\,\mathbb{E}\left(Q^{N-1} r(X_1)\Big|\mathcal{F}_0\right) \\ &= q(X_0) + \mathbb{E}\, PQ^{N-1} r(X_0) = Q^N r(X_0) = \mathbb{E}\, Q^N r(X_0). \end{aligned}$$

Jeżeli deterministyczne x_0 spełnia $Q^N r(X_0) = r(X_0)$, to $\hat\tau = 0$ oraz

$$J_{N,r}(\hat\tau, X_0) = r(X_0) = Q^N r(X_0) = \mathbb{E}\, Q^N r(X_0).$$

Jeśli x_0 jest losowe, to

$$J_{N,r}(\hat\tau, X_0) = \int_E \mathbb{E}\left(J_{N,r}(\hat\tau, X_0)\big| X_0 = x\right)\mu(dx) = \int_E J_{N,r}(\hat\tau, x)\mu(dx),$$

gdzie μ jest rozkładem X_0, co sprowadza nas do przypadku deterministycznego warunku początkowego. $\square$

16.1 Równania filtracji

Niech $X = (X_n)$ będzie łańcuchem Markowa na przestrzeni stanów $(E, \mathcal{E})$. Niech $(P_n(x, \Gamma))$, $x \in E$, $\Gamma \in \mathcal{E}$ będzie ciągiem prawdopodobieństw przejścia dla X, to znaczy

$$P_n(x, \Gamma) = \mathbb{E}\,(X_{n+1} \in \Gamma | X_n = x), \qquad n = 0, 1, \ldots,\ x \in E,\ \Gamma \in \mathcal{E}.$$

Proces obserwacji $Y_n = (Y_n)$ przyjmuje wartości w przestrzeni $(E_Y, \mathcal{E}_Y)$. Zakładamy, że zadany jest wzorem

$$Y_n = H(X_n, \eta_n), \qquad n = 0, 1, \ldots,$$

gdzie (η_n) jest ciągiem niezależnych zmiennych losowych w $(S, \mathcal{S})$ o tym samym rozkładzie, a H jest mierzalnym odwzorowaniem $E \times S$ w E_Y.

Oznaczmy teraz przez π_n prawdopodobieństwo warunkowe $\mathbb{P}(X_n|\mathcal{Y}_n)$, gdzie

$$\mathcal{Y}_n = \sigma(Y_1, \ldots, Y_n).$$

Celem jest znalezienie *równań filtracji*, czyli związku postaci

$$\pi_{n+1} = M_n(\pi_n, Y_{n+1}), \qquad n = 0, \ldots \tag{16.3}$$

z odpowiednio dobranymi przekształceniami (M_n).

Niech

$$\mathcal{L}(x, \varGamma) := \mathbb{P}\,(H(x, \eta_0) \in \varGamma), \qquad x \in E,\ \varGamma \in \mathcal{E}_Y.$$

Oczywiście

$$\mathcal{L}(X_{n+1}, \cdot) = \mathbb{P}\,(Y_{n+1}|X_{n+1}, \mathcal{Y}_n) \tag{16.4}$$

Zakładamy, że istnieje miara $\mathcal{L}$ na $(E_Y, \mathcal{E}_Y)$, która dominuje wszystkie prawdopodobieństwa $\mathcal{L}(x, \cdot)$. Tak więc zakładamy, że istnieje gęstośc $r : E \times E_Y \to [0, \infty)$, dla której

$$\mathcal{L}(x, \varGamma) = \int_\varGamma r(x, y)\mathcal{L}(dy), \qquad x \in E,\ \varGamma \in \mathcal{E}_Y.$$

Oznaczmy przez $\mathcal{P}(E)$ zbiór wszystkich miar probabilistycznych na $(E, \mathcal{E})$. Niech

$$P(\mu, \varGamma) = \int_E P(x, \varGamma)\mu(dx), \qquad \mu \in \mathcal{P}(E),\ \varGamma \in \mathcal{E}$$

oraz niech

$$\mathcal{M}_n : \mathcal{P}(E) \times E_Y \to \mathcal{P}(E)$$

dane będzie wzorem

$$\mathcal{M}_n(\mu, y)(A) = \frac{\int_A r(z, y)P_n(\mu, dz)}{\int_E r(z, y)P_n(\mu, dz)}, \qquad \mu \in \mathcal{P}(E)\ y \in E_Y,\ A \in \mathcal{E}. \tag{16.5}$$

Mamy następujący rezultat.

Twierdzenie 16.2 *Dla ciągu* (M_n) *zdefiniowanego przez* (16.5) *zachodzą równania filtracji* (16.3).

Dowód Mamy pokazać że dla dowolnych n, $A \in \mathcal{E}$, $C \in \mathcal{E}_Y$ i funkcji mierzalnej ograniczonej

$$F : E_Y \times E_Y \times \ldots \times E_Y \to \mathbb{R}$$

zachodzi

$$\begin{aligned}&\mathbb{E}\,\mathcal{M}_n(\pi_n, Y_{n+1})(A)\chi_C(Y_{n+1})F(Y_1, \ldots, Y_n)\\ &\quad= \mathbb{E}\,\chi_A(X_{n+1})\chi_C(Y_{n+1})F(Y_1, \ldots, Y_n).\end{aligned}$$

W tym celu ustalmy n, A, C i F. Wówczas

$$\begin{aligned}I &:= \mathbb{E}\,\mathcal{M}_n(\pi_n, Y_{n+1})(A)\chi_C(Y_{n+1})F(Y_1, \ldots, Y_n)\\ &= \mathbb{E}\,\mathbb{E}\,(\mathcal{M}_n(\pi_n, Y_{n+1})(A)\chi_C(Y_{n+1})F(Y_1, \ldots, Y_n)|X_{n+1}, \mathcal{Y}_n)\\ &= \mathbb{E}\,\mathbb{E}\,(\mathcal{M}_n(\pi_n, Y_{n+1})(A)\chi_C(Y_{n+1})|X_{n+1}, \mathcal{Y}_n)\,F(Y_1, \ldots, Y_n).\end{aligned}$$

Z Lematu **??** i (16.4) mamy

$$\mathbb{E}\left(\mathcal{M}_n(\pi_n, Y_{n+1})(A)\chi_C(Y_{n+1})|X_{n+1}, \mathcal{Y}_n\right)$$
$$\int_{E_Y} \mathcal{M}_n(\pi_n, y)(A)\chi_C(y)r(X_{n+1}, y)\mathcal{L}(dy)$$

Stąd

$$I = \mathbb{E}\int_{E_Y} \mathcal{M}_n(\pi_n, y)(A)\chi_C(y)r(X_{n+1}, y)\mathcal{L}(dy)F(Y_1, \ldots, Y_n)$$
$$\mathbb{E}\int_{E_Y} \mathcal{M}_n(\pi_n, y)(A)\chi_C(y)\,\mathbb{E}\left(r(X_{n+1}, y)|X_n, \mathcal{Y}_n\right)\mathcal{L}(dy)F(Y_1, \ldots, Y_n).$$

Teraz, z własności Markowa i z niezależności $\{\eta_n\}$ od X otrzymujemy

$$\begin{aligned}\mathbb{E}\left(r(X_{n+1}, y)|X_n, \mathcal{Y}_n\right) &= \mathbb{E}\left(r(X_{n+1}, y)|X_n\right)\\ &= \int_E r(z, y)P_n(X_n, dz).\end{aligned}$$

Reasumując

$$\begin{aligned}I = \mathbb{E}\int_{E_Y} &\mathcal{M}_n(\pi_n, y)(A)\chi_C(y)\\ &\times\mathbb{E}\left(\int_E r(z, y)P_n(X_n, z)|\mathcal{Y}_n\right)\mathcal{L}(dy)F(Y_1, \ldots, Y_n).\end{aligned}$$

Teraz

$$\mathbb{E}\left(\int_E r(z, y)P_n(X_n, dz)|\mathcal{Y}_n\right) = \int_E r(x, y)P_n(\pi_n, x).$$

Stąd

$$I = \mathbb{E}\int_{E_Y} \mathcal{M}_n(\pi_n, y)(A)\chi_C(y)\int_E r(z, y)P_n(\pi_n, dz)\mathcal{L}(dy)F(Y_1, \ldots, Y_n).$$

Z definicji $\mathcal{M}_n$ mamy

$$\mathcal{M}_n(\pi_n, y)(A)\int_E r(z, y)P_n(\pi_n, dz) = \int_A r(z, y)P_n(\pi_n, dz).$$

Czyli ostatecznie

$$\begin{aligned}I &= \mathbb{E}\int_{E_Y} \chi_C(y)\int_A r(z, y)P_n(\pi_n, dz)\mathcal{L}(dy)F(Y_1, \ldots, Y_n)\\ &= \mathbb{E}\int_{E_Y} \chi_C(y)\int_A r(z, y)\mathbb{E}\left(P_n(X_n, dz)|\mathcal{Y}_n\right)\mathcal{L}(dy)F(Y_1, \ldots, Y_n)\\ &= \mathbb{E}\int_{E_Y} \chi_C(y)\,\mathbb{E}\left(\int_A r(X_{n+1}, y)\middle|\,\mathcal{Y}_n, X_n\right)\mathcal{L}(dy)F(Y_1, \ldots, Y_n)\\ &= \mathbb{E}\, int_{E_Y}\chi_C(y)\,\chi_A(X_{n+})r(X_{n+1}, y)\mathcal{L}(dy)F(Y_1, \ldots, Y_n)\\ &= \mathbb{E}\,\chi_C(Y_{n+1})\,\chi_A(X_{n+1})F(Y_1, \ldots, Y_n),\end{aligned}$$

co kończy dowód. □

Ważnym wnioskiem z rwnania filtracji jest własność markowa filtru (π_n). Dokładnie mamy następujący rezultat.

Twierdzenie 16.3 *Proces* $(\pi_n, \mathcal{Y}_n)$ *jest markowa na przestrzeni stanów* $\mathcal{P}(E)$ *z rodziną operatorów przejścia*

$$\mathcal{P}_n(\mu)\Phi = \int_E \int_{E_Y} \Phi\left(\mathcal{M}_n(\mu, y)\right) r(z, y)\mathcal{L}(dy)P_n(\mu, dz),$$

$\mu \in \mathcal{P}(E)$, $\Phi \in B(\mathcal{P}(E))$.

Dowód

□

17

Skończone przestrzenie stanów i sterowań

Załóżmy, że

$$E = \{x_1, \ldots, x_d\}, \qquad U = \{u_1, \ldots, u_m\},$$

gdzie $d, m \in \mathbb{N}$. Dla $i \in \{1, \ldots, d\}$ niech

$$U(x_i) = \{u_1^i, \ldots, u_{m_i}^i\} \subset U$$

będzie zbiorem dopuszczalnych parametrów sterujących gdy proces jest w stanie x_i.

Niech $\{P^u; u \in U\}$ będzie rodziną operatorów przejścia; $P^u = [P_{ij}^u] \in M(d \times d)$,

$$P_{ij}^u = \mathbb{E}\left(X_{n+1}^\pi = x_j | u_n = u, X_n^\pi = x_i\right).$$

Oczywiście $P_{i,j}^u \in [0,1]$ oraz

$$\sum_{j=1}^d P_{i,j}^u = 1, \qquad i = 1, \ldots, d.$$

Mamy zadane koszty/zyski bieżące

$$0 \le q(x,u), \qquad x \in E,\ u \in U(x),$$

oraz dyskąto $y \in (0,1]$. Funkcjonał kosztu/zysku na nieskończonym horyzoncie czasowym dany jest wzorem

$$J(X_0, \pi) = \mathbb{E} \sum_{n=0}^{\infty} q\left(x_n^\pi, u_n(X_0^\pi, \ldots, x_n^\pi)\right).$$

Przypomnijmy, że

$$\mathcal{A}h(x) = \max_{u \in U(x)} \{q(x,u) + \gamma P^u h(x)\}$$

gdy J jest zyskiem, a

$$\mathcal{A}h(x) = \min_{u \in U(x)} \{q(x,u) + \gamma P^u h(x)\}$$

gdy J jest kosztem.

Niech

$$V_n = \mathcal{A}^n(0), \qquad V_\infty = \lim_{n\to\infty} V_n.$$

Twierdzenie 17.1 *(i)* $V_\infty \colon E \to [0,+\infty)$ *jest funkcją ograniczoną. (ii)* V_∞ *jest jedynym nieujemnym rozwiązaniem równania*

$$V(x) = \mathcal{A}V(x), \qquad x \in E.$$

(iii) Istnieje $v_\infty \colon E \to U$ *takie, że* $v_\infty(x) \in U(x)$, $x \in E$, *oraz*

$$V_\infty(x) = q(x, v_\infty(x)) + \gamma P^{v_\infty(x)} V_\infty(x), \qquad x \in E.$$

(iv) Strategia $\widehat{\pi} = (\widehat{u}_n)$, *gdzie*

$$\widehat{u}_n(x_0,\ldots,x_n) = v_\infty(x_n),$$

jest otymalna i $V_\infty(x) = J(x,\widehat{\pi})$, $x \in E$.

Dowód

18

Zadania

Zadanie 18.1 Niech $\{\xi_n\}$ będzie ciągiem iid elementów losowych określonych na przestrzeni probabilistycznej $(\Omega, \mathcal{F}, \mathbb{P})$ o wartościach w przestrzeni mierzalnej $(S, \mathcal{S})$. Niech $(E, \mathcal{E})$ i $(U, \mathcal{U})$ będą przestrzeniami mierzalnymi i niech $F : E \times U \times S \to E$ będzie odwzorowaniem mierzalnym. Załóżmy, że proces sterowany zadany jest rekurencyjnie

$$X_{n+1} = F(X_n, u_n, \xi_{n+1}), \qquad X_0 = x \in E.$$

Niech $M, K \in \mathbb{N}$, $K \leq M$, $A \in \mathcal{E}$. Niech $k_0, \ldots, k_M$, f będą funkcjami mierzalnymi na E oraz niech $a \in \mathbb{R}$.

Zmieniając przestrzeń stanów sprowadzić do postaci

$$J(\pi, Y_0) = \mathbb{E}\left\{\sum_{n=0}^{N-1} q_n(Y_n, v_n) + r_N(Y_N)\right\},$$

gdzie (Y, v) jest odpowiednio skonstruowanym sterowanym procesem Markowa oraz $N < \infty$, następujące funkcjonały

(a) $\mathbb{E} \min_{0 \leq j \leq M} f(X_j)$.
(b) $\mathbb{E} \max_{0 \leq j \leq M} f(X_j)$.
(c) $\mathbb{P}(\exists j \in \{0, 1, \ldots, M\} : \ X_j \in A)$.
(d) $\mathbb{P}(\forall j \in \{0, 1, \ldots, M\} : \ X_j \in A)$.
(e)
$$\mathbb{P}\left(\sum_{n=0}^{M-1} k_n(X_n) \geq a\right) + \mathbb{E}\, k_M(X_M).$$
(f)
$$\mathbb{P}\left(\max_{j \leq M} f(X_j) \geq a\right).$$
(g)
$$\mathbb{P}\left(X_j \in A \ \textit{dla dokładnie } K\textit{-jotów} \leq M\right).$$

(h) Funkcjonał średniego czasu dojścia do zbioru A, to zanczy

$$\mathbb{E}\,\tau_A \qquad i \qquad \mathbb{E}\,\tau_A \wedge M$$

gdzie

$$\tau_A = \inf\{n : \ X_n \in A\}.$$

Zadanie 18.2 Przypuśćmy, że handlowiec ma zapas $X_n \in [0,1]$ pewnego towaru w magazynie rankiem n-tego dnia przed otwarciem sklepu. Niech $D_n \in [0,1]$ oznacza popyt na towar w ciągu n-tego dnia oraz, przy oznaczeniu $x^+ = \max\{x, 0\}$ niech

$$u_n \in [-(X_n - D_n)^+, 1 - (X_n - D_n)^+]$$

będzie zamówioną dodatkową ilością towaru, lub gdy $u_n < 0$ zwrotem towaru dokonanymi przez handlowca wieczorem n-tego dnia, a realizowanymi w następny dzień rano przed otwarciem magazynu. O $\{D_n\}$ zakładamy, że są iid o rozkładzie jednostajnym na $[0,1]$. Na koszty ponoszone w n-tym dniu składają się: koszty wynikłe z posiadania niewystarczającego zapasu

$$a\,(D_n - X_n)^+,$$

koszty składowania przez noc bX_n oraz transportu $c|u_n|$. Sformułować problem minimalizacji kosztów na skończonym odcinku czasowym. Napisać równania Bellmana.

Zadanie 18.3 Wprowadźmy dodatkowy element do poprzedniego zadania. Załóżmy, że objętość magazynu jest skończona i wynosi $K \in (0,1)$. Możliwe są dwie interpretacje. Albo dopuszczalne sterowanie u_n spełnia

$$u_n \in [-(X_n - D_n)^+, K - (X_n - D_n)^+]$$

albo nadmiar

$$X_{n+1} - K = (X_n - D_n)^+ + u_n - K$$

musi zostać składowany u sąsiada oraz koszt składowania wynosi

$$m(X_{n+1} - K).$$

Podać wzór na koszt, napisać równania Bellmana.

Zadanie 18.4 Proces produkcyjny dzieli się na trzy stadia $1, 2, 3$. Koszt utrzymania jednostki produkowanego wyrobu w stadium i wynosi s_i za jednostkę czasu. Niech w chwili n w stadium i znajduje się M_n^i jednostek. Koszt przeprowadzenia jednostki produkowanego wyrobu ze stadium $i-1$ do stadium i, $i > 1$, wynosi c_i. Koszt doprowadzenia (z zewnątrz) do stadium 1 wynosi c_1. Przypuśćmy, że w chwili m zapotrzebowanie na produkt końcowy (stadium 3) wynosi ξ_m, przy czym $\{\xi_m\}$ jest ciągiem iid. Jeżeli zapotrzebowanie ξ_m przewyższa podaż M_3^m, wprowadzamy opłatę c za każdą brakującą jednostkę.

Przypuśćmy, że interesuje nas przebieg procesu w przedziale czasowym $[0, N]$ oraz, że istnieje koszt końcowy, wynoszący f_i za jednostkę w stadium i nie sprzedaną do chwili N. W każdej chwili $m = 0, 1, \dots, N-1$, producent musi zadecydować, ile jednostek doprowadzić do stadium i, $i = 1, 2, 3$. Doprowadzenie jednostki albo do następnego stadium (ze stadium poprzedniego), albo do stadium 1, zabiera jednostkę czasu, wszystkie zaś decyzje są funkcjami czasu i liczb jednostek pozostających (w chwili podejmowanych decyzji) w każdym ze stadiów.

Sformułować problem sterowania, zapisać równania Bellmana.

Zadanie 18.5 Niech $E = \{\dots - 3, -2, -1, 0\}$, $U = \{0, 1\}$. Mamy sterowany łańcuch Markowa. Naszym celem jest dosterowania go do punktu pochłaniającego 0. Zadane są prawdopodobieństwa przejścia

$$p^0_{i,i+1} = 0.5 = p^0_{i,i-1}, \qquad p^1_{i,i+1} = 0.75, \quad p^1_{i,i-1} = 0.25, \qquad i < 0,$$

$p^i_{0,0} = 1$, $i \in U$. Zadane są: skończony horyzont czasowy N oraz funkcjonał kosztu

$$J((u_n), X_0) = \mathbb{E}\left(\sum_{n=0}^{N-1} u_n + X_N^2\right).$$

Znaleźć optymalną strategię.

Zadanie 18.6 Niech $\{Z_n\}$ będzie sterowanym łancuchem Markowa na $\mathbb{R}$ zadanym rekurencyjnie

$$Z_{n+1} = \alpha Z_n + u_n + \xi_{n+1}.$$

Zakładamy, że $\{\eta_n\}$ jest ciągiem niezależnych zmiennych losowych o jednakowym rozkładzie na $\mathbb{R}$. Celem jest jak najbliższe dosterowanie Z do procesu

$$C_{n+1} = \beta C_n + \eta_{n+1}.$$

Tutaj $\{\eta_n\}$ jest ciągiem iid na $\mathbb{R}$ niezależnym od $\{\xi_n\}$. Sterowanie u_n może zależeć od obserwowalnych wielkości Z_n i C_n.

Koszty związane ze sterowaniem wynoszą

$$J((u_n), Z_0, C_0) = \mathbb{E}\left\{\sum_{n=0}^{N-1} q(u_n) + h(Z_N - C_N)\right\}.$$

Sformułować problem sterowania, zapisać równania Bellmana. Znaleźć optymalną strategię dla $q(u) = u^2 = h(x)$, $\alpha = 1 = -\beta$, przy założeniu, że ξ_n i η_n mają średnie 0.

Zadanie 18.7 Niech $M > 0$ będzie objętością zbiornika wody, X_n ilością wody w zbiorniku w chwili n, $u_n \in [0, X_n]$ kontrolowanym wypływem wody, a ξ_n losową ilością wody wpływającej do zbiornika w chwili n. Zakładamy, że $\{\xi_n\}$ jest ciągiem iid o rozkładzie jednostajnym na $[0, M]$. Dynamika procesu $X = (X_n)$ dana jest wzorem

$$X_{n+1} = \min\{X_n - u_n + \xi_{n+1}, M\}.$$

Naszym celem jest utrzymanie wody poniżej ustalonego poziomu $L < M$. Funkcjonał kosztu uwzględnia koszty odprowadzenie wody u_n i koszty przekroczenia poziomu L. Dokładnie jest on postaci

$$J((u_n), X_0) = \mathbb{E}\left\{\sum_{n=0}^{N-1}(q(u_n) + h(X_n)) + h(X_N)\right\}$$

gdzie q jest funkcją mierzalną nieujemną i ograniczoną,

$$h(x) = a\,(x - L)^+,$$

a a jest ustalonym parametrem. Napisać równania Bellmana. Znaleźć optymalną strategie dla $q(u) = u$.

Zadanie 18.8 Niech popyt w chwili n na dany produkt przedsiębiorstwa jest modelowany łańcuchem Markowa $Y = (Y_n)$ na przestrzeni stanów $[0, \infty)$. Dynamika produkcji tego produkt dana jest wzorem

$$Z_{n+1} = Z_n + u_n,$$

tutaj u_n jest wzrostem wydajności w chwili n. Zakładamy, że wzrost wydajności u_n nie przekracza ustalonej liczby M. Ze wzrostem wydajności u_n związane są koszty bu_n. Zysk przedsiębiorstwa w chwili n jest równy dochodowi ze sprzedaży pomniejszonemu o koszty poniesione na zwiększenie wydajności, czyli jest równy

$$\lambda^n\left(\min\{Y_n, Z_n\} - bu_n\right),$$

$\lambda \in (0, 1]$ i $b > 0$ są zadanymi parametrami. Rozwiązać problem maksymalizacji zysku na skończonym i nieskończonym przedziale czasowym gdy $\{Y_n\}$ jest ciągiem iid o rozkładzie wykładniczym z zadanym parametrem γ.

Zadanie 18.9 Mamy M bezpieczników . Czas życia każdego bezpiecznika jest losowy i wynosi j z prawdopodobieństwem p_j, oczywiście $\sum p_j = 1$. Zaplanowana wymiana jednego bezpiecznika kosztuje a. Awaria bezpiecznika kosztuje $b >> a$. Koszt związany z awarią nie zależy od liczby bezpieczników, które uległy jednoczesnemu zepsuciu. Z drugiej strony wadliwe bezpieczniki muszą być wymienione i koszt (dodatkowy do kosztu awarii) wymiany m bezpieczników wynosi ma.

Sprowadzić do postaci standartowej zagadnienie znalezienia optymalnej strategi wymiany bezpieczników na skończonym i nieskończonym przedziale czasowym. Uwzględnić dyskonto $\lambda \in (0, 1]$. Rozważyć problem minimalizacji kosztów na jednostkę czasu.

Zadanie 18.10 Niech $\{Y_n\}$ będzie kontrolowanym gałązkowym procesem rozpadu neutronów. W szczególności niech $p_i(\alpha)$, $i = 0, 1, 2, \ldots$, $\alpha \in \mathbb{R}$, oznacza prawdopodobieństwo zdarzenia, że dany neutron produkuje w jednostkowym odcinku czasowym i neutronów przy parametrze sterującym α. Naszym celem jest uzyskanie M neutronów w najkrótszym czasie, przy czym chcemy uniknąć groźnej sytuacji gdy

$$Y_n \geq \frac{3M}{2}.$$

Parametr α zmieniony może być tylko pośrednio, to znaczy

$$\alpha_{n+1} = \alpha_n + \delta, \qquad \delta = -1, 0, 1.$$

Niech koszt ma postać

$$J((\pi), Y_0) = \mathbb{E}\tau + c\mathbb{P}\left(\sup_{0 \leq n \leq \tau} Y_n \geq \frac{3M}{2}\right),$$

gdzie τ jest pierwszym momentem, w którym $Y_n \geq M$. Sprowadzić problem do postaci standartowej oraz zapisać równanie funkcyjne dla kosztu minimalnego. Wobec istnienia pamięci urządzenia sterującego niezbędna jest zmiana przestrzeni stanów.

Zadanie 18.11 Niech $\{(Y_n, Z_n)\}$ będzie procesem Markowa o wartościach w $[0, 1] \times [0, 1]$. Interpretujemy Y i Z jako odsetki uzyskane w chwili n z kapitału ulokowanego w przedsięwzięciach 1 i 2. Sterowanie polega na decydowaniu, w którym przedsięwzięciu ulokować pieniądze. Niech X_n oznacza kapitał inwestora w chwili n. Lokata w przedsięwzięciu 1 daje w chwili $n+1$, $X_n(1+Y_n)$ a w przedsięwzięciu 2, $X_n(1 + Z_n)$. Inwestowanie ustaje gdy X_n przekracza ustalony poziom x. Koszt równa się średniemu czasowi potrzebnemu na osiągnięcie poziomu x. Istnieje przy tym koszt $c > 0$ związany z przeniesieniem kapitału z jednego przdsięwzięcia do drugiego. Sprowadzić problem do postaci standartowej.

Zadanie 18.12 Przypuśćmy, że mamy dwie skrzynki: A i B; w jednej z nich znajduje się piłka. W przypadku gdy zaglądamy do skrzynki A i piłka znajduje się w tej właśnie skrzynce, prawdopodobieństwo jej zauważenia wynosi $p_1 \in (0, 1)$. Podobnie, prawdopodobieństwo zauważenia piłki w skrzynce B (jeżeli się tam znajduje) wynosi $p_2 \in (0, 1)$. Niech x oznacza prawdopodobieństwo a priori tego, że piłka znajduje się w skrzynce A. Zadanie polega na wyznaczeniu strategii szukania minimalizującej średni czas potrzebny na zlokalizowanie piłki, to znaczy należy określić regułe decyzyjną wskazującą najlepszą skrzynkę, do której należy zajrzeć w chwili n, przy danej historii poprzednich sprawdzeń zawartości skrzynek. Określić stosowną przestrzeń stanów oraz proces Markowa opisujący problem. Napisać równania Bellmana. Pokazać, że istnieje przynajmniej jedna strategia zapewniająca znalezienie piłki w skończonym czasie. (Zbadaj strategię zmieniania za każdym razem sprawdzanej skrzynki).

Zadanie 18.13 Rozważyć wariant poprzedniego zadania z m skrzynkami i prawdopodobieństwami zobaczenia piłki $p_i \in (0,1)$.

Zadanie 18.14 Rozważyć wariant Zadania 18.12, gdzie piłka zmienia za każdym razem skrzynkę z prawdopodobieństwem $q \in (0,1)$.

Zadanie 18.15 Rozważmy wariant poprzedniego zadania (i Zadania 18.13) uwzględniający średni koszt na jednostkę czasu. Przypuśćmy, że koszt jednostkowy związany jest z popełnieniem błędu (piłka nie zostaje znaleziona w sprawdzanej skrzynce), z sukcesem natomiast (piłka zostanie znaleziona w sprawdzanej skrzynce) związany jest koszt zerowy. Sformułować problem średniego kosztu na jednostkę czasu i zapisać dokładne równanie funkcyjne dla kosztu optymalnego.

Zadanie 18.16 Prezydent Stanów Zjednoczonych polecił dyrektorowi FBI ujęcie groźnego terrorysty. Z badań operacyjnych wiadomo, że może on się ukrywać w m różnych miastach. Każdej nocy zmienia on miejsce ukrycia zgodnie z macierzą prawdopodobieństw przejścia

$$P = (p_{i,j}), \quad i, j = 1, \ldots, m.$$

Tutaj $p_{i,j}$ to prawdopodobieństwo warunkowe, że terrorysta będąc w i-tym mieście przemieści się w ciągu nocy do miasta j-tego.

Dostępne środki pozwalają na jednoczesne przeszukanie co najwyżej k - miast. Oczywiście $k < m$. Niestety nawet wtedy efektywność działań nie jest stuprocentowa. Jeżeli poszukiwany terrorysta znajduje sie w j-tym mieście to prawdopodobieństwo jego znalezienia wynosi $q_j \in (0,1)$. Wiadomy jest rozkład początkowy prawdopodobieństw $v = (v_1, \ldots, v_m)$ pobytu terrorysty. Pomóż dyrektorowi znaleźć strategie poszukiwań minimalizującą oczekiwany czas pochwycenia terrorysty.

Zadanie 18.17 Przypuśćmy, że dystrybutor może łatwo psujący się towar umieścić w dwóch punktach sprzedaży, oznaczonych numerami 1 i 2. Towar, który nie zostaje sprzedany w dniu jego wyprodukowania, psuje się i nie może być już sprzedany. Niech ciągi zapotrzebowań (popytów) w obydwu punktach sprzedaży mają postać $\{d_n^1\}$ i $\{d_n^2\}$, gdzie wszystkie zmienne losowe są wzajemnie niezależne. Zmienne $\{d_n^1\}$ mają te same rozkłady podobnie jak zmienne $\{d_n^2\}$. Dzienna produkcja wynosi m jednostek towaru. Ile jednostek powinno się znaleźć w każdym z punktów sprzedaży, aby średnia liczba jednostek nie sprzedanych była minimalna?

Zadanie 18.18 Rozważyć następujący wariant poprzedniego zadania: towary nie ulegają zepsuciu, to znaczy towary nie sprzedane jednego dnia mogą być sprzedane któregoś innego dnia, bez zmiany ceny. Jaką postać ma odpowiednia przestrzeń stanów oraz proces Markowa opisujący problem? Zapisać dokładnie równania funkcyjne dla problemu "minimalnej średniej liczby towarów nie sprzedanych w ciągu dnia".

Zadanie 18.19 Powtórzyć dwa poprzednie zadania przy założeniu, że ciągi zapotrzebowań są procesami Markowa postaci

$$d^i_{n+1} = a_i d^i_n + \xi^i_n,$$

gdzie $\{\xi^i_n\}$ są wzajemnie niezależnymi ciągami iid.

Zadanie 18.20 Niech X_n oznacza wagę hodowanej przez rolnika świni w n-tym dniu. Każdego dnia rolnik decyduje czy zwierze hodować dalej czy też go sprzedać, uzyskując cene $c(X_n)$, gdzie c jest rosnącą funkcją spełniającą $c(0) = 0$. Dzienny koszt utrzymania zwierzęcia o wadze x wynosi $a(x)$ gdzie a dana funkcja nieujemna niemalejąca. Waga X_n jest procesem Markowa ze stanem pochłaniającym 0, który interpretujemy jako stan w którym zwierze zachorowało i nie będzie nadawać się do spożycia. Sformułować problem optymalnego wyboru chwili sprzedaży (wyboru losowego momentu sprzedaży) jako problem optymalnego zatrzymania. Rozpatrzyć przypadek gdy operator przejście P dla X jest postaci

$$P\psi(x) = p\psi(x+1) + (1-p)\psi(0).$$

Zinterpretować taki proces.

Zadanie 18.21 Przyjmijmy w poprzednim zadaniu, że stan świni nie jest dokładnie znany lecz jej mechanizm wzrostu pozostaje niezmieniony. Świnia rośnie aż do momentu zachorowania, kiedy staje się bezwartościowa. W każdej chwili n przeprowadza się test pomagający oszacować jej stan zdrowia. Wynik testu ξ_n (o wartościach 1 lub 0) ma następujący rozkład: jeżeli jest ona jest zdrowa, zmienna losowa ξ_n ma rozkład $p_z(\xi)$, jeśli natomiast jest chora - rozkład $p_0(\xi)$. Sformułować problem optymalnego zatrzymania.
Uwaga. Przy każdej z hipotez dotyczących stanu zdrowia świni zmienne losowe ξ_n są wzajemnie niezależne. Niech $Y_n = \mathbb{P}(Z|\xi_0, \ldots \xi_{n-1})$. Zauważmy, że

$$X_n = \begin{cases} X_n + n, & \text{z prawdopodobieństwem } Y_n, \\ 0, & \text{z prawdopodobieństwem } 1 - Y_n. \end{cases}$$

Zadanie 18.22 Niech

$$\begin{aligned} Y_{n+1} &= f_y(Y_n, u(Y_n, Z_n), \xi_n), \\ Z_{n+1} &= f_z(Z_n, \psi_n) \end{aligned}$$

będą procesami Markowa, gdzie $\{\xi_n\}$ i $\psi_n\}$ są niezależnymi ciągami iid. Proces Y reprezentuje położenie obiektu fotografującego, którego zadaniem jest zbliżenie się do obiektu fotografowanego o położeniu Z_n, tak by możliwe było zrobienie dobrego zdjęcia. Możliwe jest wykonanie tylko jednego zdjęcia - wybrać należy właściwy moment. Kosztem bieżącym jest koszt paliwa, równy w chwili n, $k(u(Y_n, Z_n))$. Wartość fotografii wynosi $g(|Y_n - Z_n|)$. Sformułować problem jako problem optymalnego zatrzymania.

Zadanie 18.23 Niech (X_n^i), $i = 1, \ldots, M$ oznacza rozmiar i-tego kłosa zboża na polu w n-tym dniu. Zakładamy, że kłosy rozwijają się niezależnie od siebie. Ponadto zakładamy, że (X_n^i), $i = 1, \ldots M$ są jednorodnymi procesami Markowa o tym samym rozkładzie. Przyjmujemy, że stan 0 jest pochłaniający dla (X_n^i). Interpretujemy $X_n^i = 0$ jako stan, w którym i-ty kłos uległ zepsuciu. Rolnik obserwuje stan kłosów i na tej podstawie podejmuje decyzje o dniu $\tau \leq N$ żniw. Jego zysk wynosi

$$J(\tau, X_0) = \mathbb{E} \sum_{i=1}^{M} X_\tau^i.$$

Rozważyć dwa warianty: 1) Wariant mało realistyczny gdzie rolnik zna dokładnie stan wszystkich kłosów. 2) Wariant realistyczny, gdzie rolnik zna dokładnie stan próbki $M_0 << M$ kłosów.

Rozważyć szczególny przypadek gdzie operator przejścia P procesu (X_n^i) jest następującej postaci

$$Pf(x) = pf(x+h) + (1-p)f(0), \qquad x \in [0, \infty),$$

gdzie $p \in (0,1)$ i $h > 0$ są ustalonymi liczbami.

19

Rozwiązania

Zadanie 18.1 (a). Niech

$$Z_n = \min\{f(X_j):\ j \le n\}, \qquad n = 0, 1, \ldots .$$

Wówczas $Y_n = (X_n, Z_n)$ jest jednorodnym procesem Markowa bo dynamikę Y_n można opisać następująco $Z_0 = f(X_0)$ i

$$Z_{n+1} = \min\{Z_n, f(X_{n+1})\} = \min\{Z_n, f(F(X_n, u_n, \xi_{n+1}))\}.$$

Oczywiście

$$\mathbb{E} \min_{0 \le j \le M} f(X_j) = \mathbb{E}\, Z_M.$$

Stąd

$$J(\pi, Y_0) = \mathbb{E}\, r(Y_M),$$

gdzie $r_M(x, z) = z$.
Zadanie 18.1 (b). Stosujemy rozumowanie z części (a), kładąc

$$Z_n = \max\{f(X_j):\ j \le n\}, \qquad n \ge 0.$$

Zadanie 18.1 (c). Ponieważ

$$\mathbb{P}\,(\exists\, j \in \{0, 1, \ldots, M\}:\ X_j \in A) = \mathbb{E} \max_{0 \le j \le M} \chi_A(X_j),$$

problem jest szczególnym przypadkiem problemu (b).
Zadanie 18.1 (d). Ponieważ

$$\mathbb{P}\,(\forall\, j \in \{0, 1, \ldots, M\}:\ X_j \in A) = \mathbb{E} \min_{0 \le j \le M} \chi_A(X_j),$$

problem jest szczególnym przypadkiem problemu (a).
Zadanie 18.1 (e). Zdefiniujmy

$$Z_n = \sum_{j=0}^{n} k_j(X_j), \qquad n \ge 0.$$

Ponieważ

$$Z_{n+1} = Z_n + k_{n+1}(X_{n+1}) = Z_n + k_{n+1}(F(X_n, u_n, \xi_{n+1})),$$

Y jest procesem Markowa. Ponadto

$$\begin{aligned}\mathbb{P}\left(\sum_{j=0}^{M-1} k_j(X_j) \geq a\right) &= \mathbb{P}\left(Z_{M-1} \geq a\right) \\ &= \mathbb{E}\chi_{[a,\infty)}(Z_{M-1}).\end{aligned}$$

Zadanie 18.1 (f). Ponieważ

$$\begin{aligned}\mathbb{P}\left(\max_{0\leq j\leq M} f(X_j) \geq a\right) &= \mathbb{P}\left(\exists j \in \{0,1,\ldots,M\}:\ X_j \in f^{-1}([a,\infty))\right) \\ &= \mathbb{E}\max_{0\leq j\leq M} \chi_{f^{-1}([a,\infty))}(X_j),\end{aligned}$$

problem sprowadza się do problemów (c) i (b).
Zadanie 18.1 (g). Ponieważ

$$\mathbb{P}\left(X_j \in A \text{ dla dokładnie } K\text{-jotów} \leq M\right) = \mathbb{P}\left(\sum_{j=0}^{M} \chi_A(X_j) = K\right)$$

problem sprowadza się do problemu (e).
Zadanie 18.1 (h). Niech $Z_0 = \chi_A(X_0)$ oraz

$$Z_{n+1} := \max\{\chi_A(X_{n+1}), Z_n\} = \max\{\chi_A(F(X_n, u_n, \xi_{n+1}), Z_n\}.$$

Oczywiście $Y_n = (Z_n, X_n)$ jest procesem Markowa. Ponadto

$$\mathbb{E}\,\tau_A = \mathbb{E}\sum_{n=0}^{\infty}(1 - Z_n)$$

oraz

$$\mathbb{E}\,\tau_A \wedge M = \mathbb{E}\sum_{n=0}^{M}(1 - Z_n).$$

Zadanie 18.2. Przy oznaczeniu $x^+ = \max\{x, 0\}$, dynamika procesu $Z_n = (X_n, D_n)$ dana jest wzorem

$$\begin{aligned}&D_{n+1} = \xi_{n+1}, \\ &X_{n+1} = (X_n - D_n)^+ + u_n(X_n, D_n),\end{aligned}$$

gdzie $\{\xi_n\}$ jest ciągiem niezależnych zmiennych losowych o rozkładzie jednostajnym na $[0,1]$, przestrzeń stanów E wynosi $[0,1]\times[0,1]$, $U = [-1,1]$, a dla pary $(x,d) \in E$ zbiór sterowań dopuszczalnych jest równy

$$U(x,d) = [-(x-d)^+, 1-(x-d)^+].$$

Funkcjonał kosztów jest postaci

$$J(\pi, X_0, D_0) = \mathbb{E}\left\{\sum_{n=0}^{N-1} q(X_n, D_n, u_n) + r(X_N, D_N)\right\},$$

gdzie

$$\begin{aligned} q(x,d,u) &= r(x,d) + c|u|, \\ r(x,d) &= a(d-x)^+ + bx. \end{aligned}$$

Operator przejścia jest postaci

$$P^u f(x,d) = \int_0^1 f\left((x-d)^+ + u, \xi\right) d\xi, \qquad u \in [-(x-d)^+, 1-(x-d)^+].$$

Operator $\mathcal{A}$ występujący w równaniach Bellmana wynosi

$$\begin{aligned} \mathcal{A}v(x,d) &= \inf_{u\in U(x,d)} \left(q(x,d,u) + P^u v(x,d)\right) \\ &= r(x,d) + \inf_{u\in U(x,d)} \left(c|u| + \int_0^1 v\left((x-d)^+ + u, \xi\right) d\xi\right). \end{aligned}$$

Mamy

$$\begin{aligned} P^u r(x,d) &= \int_0^1 r\left((x-d)^+ + u, \xi\right) d\xi \\ &= \int_0^1 \left\{a(\xi - (x-d)^+ - u)^+ + b(x-d)^+ + bu\right\} d\xi \\ &= \frac{a}{2}\left(1-(x-d)^+ - u\right)^2 + b(x-d)^+ + bu. \end{aligned}$$

Tak więc $\mathcal{A}r(x,d)$ wynosi

$$r(x,d) + b(x-d)^+ + \inf_{u\in U(x,d)} \left(c|u| + bu + \frac{a}{2}\left(1-(x-d)^+ - u\right)^2\right).$$

Dalsze rachunki pozostawiamy czytelnikowi.

Zadanie 18.3 W pierwszej wersji zbiór sterowań dopuszczalnych wynosi

$$U(x,d) = [-(x-d)^+, K-(x-d)^+].$$

W drugiej wersji $U(x,d)$ jest taki sam jak w Zadaniu 18.2 ale koszt jest postaci

$$J(\pi, X_0, D_0) = \mathbb{E}\left\{\sum_{n=0}^{N-1} \tilde{q}(X_n, D_n, u_n) + r(X_N, D_N)\right\}$$

gdzie
$$\tilde{q}(x,d,u) = q(x,d,u) + m\left((x-d)^+ + u - K\right)^+,$$
a q występuje w funkcjonale kosztu z poprzednim zadaniu.
Zadanie 18.4 Jako przestrzeń stanów przyjmijmy
$$E = \mathbb{N}\times\mathbb{N}\times\mathbb{N}\times\mathbb{N}.$$
Przyjmujemy $\mathbb{N} = \{0,1,\ldots\}$. Przestrzenią sterowań jest
$$U = \mathbb{N}\times\mathbb{N}\times\mathbb{N}.$$
Przyjmijmy, że
$$X_n = (M_n^1, M_n^2, M_n^3, D_n).$$
Dynamika X jest następująca
$$\begin{aligned}
M_{n+1}^1 &= M_n^1 - u_n^2 + u_n^1,\\
M_{n+1}^2 &= M_n^2 - u_n^3 + u_n^2,\\
M_{n+1}^3 &= \left(M_n^3 + u_n^3 - D_n\right)^+ = \max\{M_n^3 + u_n^3 - D_n, 0\},\\
D_{n+1} &= \xi_{n+1}.
\end{aligned}$$
Tak więc sterowanie u^1 reprezentuje import produktu do stanu 1, a sterowania u^2 i u^3 reprezentują transfer produktu ze stanu 1 do stanu 2 i ze stanu 2 do końcowego stanu 3. Mamy następujące ograniczenia na sterowania
$$\begin{aligned}
M^1 - u^2 + u^1 &\geq 0,\\
M^2 - u^3 + u^2 &\geq 0.
\end{aligned}$$
W chwili n koszt wynosi
$$q(X_n, u_n) = c_1 u_n^1 + c_2 u_n^2 + c_3 u_n^3 + c\left(D_n - M_n^3\right)^+.$$
Koszt końcowy wynosi
$$r(X_N) = f_1 M_N^1 + f_2 M_N^2 + f_3 M_N^3.$$
Ostatecznie funkcjonał kosztu na przedziale $[0,N]$ jest postaci
$$J_N(\pi, X_0) = \mathbb{E}\left\{\sum_{n=0}^{N-1} q(X_n, u_n) + r(X_N)\right\}.$$
Zadanie 18.5 Mamy
$$q(x,u) = u, \qquad r(x) = x^2.$$
Ponadto

$$P^0 v(x) = \begin{cases} \frac{1}{2}\left(v(x-1)+v(x+1)\right) & \text{gdy } x<0, \\ v(0) & \text{gdy } x=0 \end{cases}$$

oraz

$$P^1 v(x) = \begin{cases} \frac{1}{4}v(x-1)+\frac{3}{4}v(x+1) & \text{gdy } x<0, \\ v(0) & \text{gdy } x=0. \end{cases}$$

Stąd, dla $x<0$,

$$\mathcal{A}v(x) = \min\left\{\frac{1}{2}\left(v(x-1)+v(x+1)\right), 1+\frac{1}{4}v(x-1)+\frac{3}{4}v(x+1)\right\}.$$

Czyli, dla $x<0$,

$$\mathcal{A}r(x) = \min\{x^2+1, x^2+x+2\} = x^2+x+2.$$

Oczywiście $\mathcal{A}v(0)=v(0)$. Stąd

$$\mathcal{A}r(x) = \begin{cases} x^2+x+2 & \text{dla } x<0, \\ 0 & \text{dla } x=0 \end{cases}$$

oraz sterowanie u_{N-1} dane jest wzorem

$$u_{N-1}(x) = \begin{cases} 1 & \text{dla } x<0, \\ 0 & \text{dla } x=0. \end{cases}$$

Oczywiście mamy $\mathcal{A}r(0)=0=\mathcal{A}^j r(0)$. Następnie, dla $x<-1$ mamy

$$\mathcal{A}^2 r(x) = \min\left\{x^2+1+2, x^2+2x+4-\frac{1}{4}+\frac{3}{4}\right\} = x^2+2x+4+\frac{1}{2}.$$

Natomiast dla $x=-1$,

$$\begin{aligned} \mathcal{A}^2 r(x) &= \min\left\{\frac{1}{2}(x-1)^2+\frac{1}{2}(x-1)+2, \frac{1}{4}(x-1)^2+\frac{1}{4}(x-1)+3\right\} \\ &= \frac{1}{2}(x-1)^2+\frac{1}{2}(x-1)+2. \end{aligned}$$

Stąd

$$u_{N-2}(x) = \begin{cases} 0 & \text{dla } x<-1, \\ 1 & \text{dla } x=-1,0. \end{cases}$$

Dalej pokażemy przez indukcje, że

$$\mathcal{A}^n r(x) = \begin{cases} x^2+nx+a_n & \text{dla } x<-n, \\ b_x^n & \text{dla } x\in[-n,0], \end{cases}$$

oraz

$$u_{N-n}(x) = \begin{cases} 1 & \text{dla } x < -n, \\ 0 & \text{dla } x \in [-n, 0], \end{cases}$$

gdzie ciągi a_n i $\{b_x^n\}$, $x \in [-n, 0]$ zadane są indukcyjnie

$$a_{n+1} = a_n + 2 + \frac{n}{2}, \qquad a_0 = 0,$$

$b_0^n = 0$,

$$b_x^{n+1} = \begin{cases} \frac{1}{2} b_{x-1}^n + \frac{1}{2} b_{x+1}^n & \text{gdy } x \in (-n, -1), \\ \frac{1}{2} b_{-2}^n & \text{gdy } x = -1, \\ \frac{1}{2}\big((-n-1)^2 - n(-n-1) + a_n\big) + \frac{1}{2} b_{-n+1}^n & \text{gdy } x = -n, \\ \frac{1}{2}\big((-n-2)^2 - n(-n-2) + a_n\big) + \frac{1}{2} b_{-n}^n & \text{gdy } x = -n-1. \end{cases}$$

Najpierw pokazujemy tezę dla ciągu $\{a_n\}$. Dla $n = 1$ pokazaliśmy, że $a_1 = 2$ co jest zgodne z podaną formułą. Załóżmy, że teza jest prawdziwa dla danego n. Wówczas, ponieważ

$$\mathcal{A}^n r(x) = x^2 + nx + a_n, \qquad x < -n$$

dla $x < -n - 1$ mamy

$$\begin{aligned} &\mathcal{A}^{n+1} r(x) \\ &= \min\left\{\frac{\mathcal{A}^n r(x-1) + \mathcal{A}^n r(x+1)}{2}, \frac{\mathcal{A}^n r(x-1) + 3\mathcal{A}^n r(x+1) + 4}{4}\right\} \\ &= \min\left\{x^2 + 1 + nx + a_n, x^2 + 1 + x + nx + \frac{n}{2} + a_n + 1\right\}. \end{aligned}$$

Wystarczy więc zauważyć, że dla $x < -n - 1$,

$$x^2 + 1 + nx + a_n > x^2 + 1 + x + nx + \frac{n}{2} + a_n + 1.$$

Pokazujemy teraz tezę dla ciągu $\{b_x^n\}$. W tym celu wystarczy pokazać jedynie pokazać, że dla $x \in [-n, 0]$, w $\mathcal{A}\mathcal{A}^{n-1} r(x)$ minimum jest osiągalne dla stategii $u = 0$. Inaczej, że zachodzą

$$\frac{1}{2}(b_{x-1}^n + b_{x+1}^n) \le \frac{1}{4}\left(b_{x-1}^n + 3b_{x+1}^n + 4\right), \quad x \in (-n, -1),$$

$$\frac{1}{2} b_{-2}^n \le \frac{1}{4}\left(b_{-2}^n + 4\right),$$

$$\begin{aligned} &\frac{1}{2}\big((-n-1)^2 - n(-n-1) + a_n\big) + \frac{1}{2} b_{-n+1}^n \\ &\le \frac{1}{4}\big((-n-1)^2 - n(-n-1) + a_n\big) + \frac{3}{4} b_{-n+1}^n + 1, \end{aligned}$$

$$\frac{1}{2}\left((-n-2)^2 - n(-n-2) + a_n\right) + \frac{1}{2}b^n_{-n}$$
$$\leq \frac{1}{4}\left((-n-2)^2 - n(-n-2) + a_n\right) + \frac{3}{4}b^n_{-n} + 1.$$

Rachunek pozostawiamy czytelnikowi.

Zadanie 18.6 Przestrzenią stanów jest $E = \mathbb{R} \times \mathbb{R}$ a przestrzenią sterowań jest $U = \mathbb{R}$. Operator przejścia dany jest wzorem

$$\begin{aligned} P^u v(z,c) &= \mathbb{E}v(\alpha z + u + \xi_0, \beta c + \eta_0) \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} v(\alpha z + u + x, \beta c + y)\psi_\xi(dx)\psi_\eta(dy), \end{aligned}$$

gdzie ψ_ξ i ψ_η są rozkładami ξ_0 i η_0. Operator Bellmana ma postać

$$\mathcal{A}v(z,c) = \inf_{u\in\mathbb{R}} \left\{q(u) + P^u v(z,c)\right\}.$$

Rozważmy przypadek gdy $q(u) = u^2$ i $r(z,c) = h(z-c) = (z-c)^2$ i $\alpha = 1 = -\beta$. Wtedy, ze scentrowania i niezależności ξ_0 i η_0 mamy

$$\begin{aligned} \mathcal{A}r(z,c) &= \inf_{u\in\mathbb{R}} \left\{u^2 + \mathbb{E}\left((z+u+c) + (\xi_0 - \eta_0)\right)^2\right\} \\ &= \mathbb{E}(\xi_0^2 + \eta_0^2) + \inf_{u\in\mathbb{R}} \left\{u^2 + (z+c+u)^2\right\} \\ &= \mathbb{E}(\xi_0^2 + \eta_0^2) + \frac{(z+c)^2}{2}, \end{aligned}$$

przy czym minimum jest osiągane dla

$$u_{N-1} = -\frac{z+c}{2}.$$

Następnie

$$\begin{aligned} \mathcal{A}^2 r(z,c) &= \frac{3}{2}\,\mathbb{E}\left(\xi_0^2 + \eta_0^2\right) + \inf_{u\in\mathbb{R}} \left\{u^2 + \frac{(z+u-c)^2}{2}\right\} \\ &= \frac{3}{2}\,\mathbb{E}\left(\xi_0^2 + \eta_0^2\right) + \frac{11(z+c)^2}{9}, \end{aligned}$$

przy czym minimum jest osiągane dla

$$u_{N-2} = -\frac{z-c}{3}.$$

Stąd

$$\mathcal{A}^n r(z,c) = a_n\,\mathbb{E}\left(\xi_0^2 + \eta_0^2\right) + \frac{\left(z + (-1)^{n+1}c\right)^2}{b_n}$$

oraz optymalną strategią na przedziale skończonym $[0,N]$ jest

$$u_n = -\frac{z + (-1)^{n+1}c}{f_{N-n}}, \qquad n = 0, \ldots, N-1,$$

gdzie, $a_0 = 1 = b_0$,

$$a_{n+1} = a_n + \frac{1}{b_n}, \qquad b_{n+1} = \frac{b_n^2 \left(1 + \frac{1}{b_n}\right)^2}{\left(b_n \left(1 + \frac{1}{b_n}\right) - 1\right)^2 + 1}$$

oraz

$$f_n = b_{N-n} \left(1 + \frac{1}{b_{N-n}}\right).$$

Zadanie 18.7 Operator Bellmana ma postać

$$\mathcal{A}v(x) = \inf_{u \in [0,x]} \left\{ q(u) + h(x) + \frac{1}{M} \int_0^M v\left(\min\{x - u + y, M\}\right) dy \right\}.$$

W naszym przypadku $r = h$. Stąd

$$\begin{aligned} \mathcal{A}h(x) = \inf_{u \in [0,x]} \Bigg\{ & q(u) + h(x) \\ & + \frac{a}{M} \int_0^M \max\{\min\{x - u + y, M\} - L, 0\} dy \Bigg\}. \end{aligned}$$

Ponieważ

$$\begin{aligned} & \int_0^M \max\{\min\{x - u + y, M\} - L, 0\} dy \\ & = \int_{L-x+u}^{M-x+u} (x - u + y - L) dy + \int_{M-x+u}^{M} (M - L) dy \\ & = \int_0^{M-L} y dy + (M - L)(x - u) \\ & = \frac{(M-L)^2}{2} + (M - L)(x - u), \end{aligned}$$

to, kładąc $\kappa = a(M - L)/M$ otrzymujemy

$$\begin{aligned} \mathcal{A}h(x) &= h(x) + \frac{\kappa(M-L)}{2} + \inf_{u \in [0,x]} \{q(u) + \kappa(x - u)\} \\ &= h(x) + \kappa x + \frac{\kappa(M-L)}{2} + \inf_{u \in [0,x]} \{q(u) - \kappa u\}. \end{aligned}$$

Dla $q(u) = u$ mamy

$$\mathcal{A}h(x) = \begin{cases} h(x) + x + \frac{\kappa(M-L)}{2} & \text{gdy } 1-\kappa \leq 0, \\ h(x) + \kappa x + \frac{\kappa(M-L)}{2} & \text{gdy } 1-\kappa > 0 \end{cases}$$

oraz

$$u_{N-1}(x) = \begin{cases} x & \text{gdy } 1-\kappa \leq 0, \\ 0 & \text{gdy } 1-\kappa > 0. \end{cases}$$

Aby policzyć dalsze iteracje $\mathcal{A}^n h$ policzymy $P^u v(x)$ dla $v(x) := x$.

$$\begin{aligned} P^u v(x) &= \frac{1}{M}\int_0^M \min\{x-u+y, M\}dy \\ &= \frac{1}{M}\left\{\int_0^{M-x+u}(x-u+y)dy + \int_{M-x+u}^M M dy\right\} \\ &= \frac{1}{M}\left\{\int_{x-u}^M z dz + M(x-u)\right\} \\ &= \frac{1}{M}\left\{\frac{1}{2}M - \frac{1}{2}(x-u)^2 + M(x-u)\right\}. \end{aligned}$$

Korzystając z powyższej równości i z rachunków przeprowadzonych dla $\mathcal{A}h$ otrzymujemy,

$$\mathcal{A}^2 h(x) = h(x) + \kappa x + \kappa(M-L) + C(x),$$

gdzie, gdy $1-\kappa \leq 0$, to

$$\begin{aligned} C(x) &= \frac{M}{2} + \inf_{u\in[0,x]}\left\{(1-\kappa)\,u + (x-u) - \frac{1}{2M}(x-u)^2\right\} \\ &= (1-\kappa)\,x + \frac{M}{2}, \end{aligned}$$

przy czym infimum jest osiągane dla $u = x$. Stąd gdy $1-\kappa \leq 0$, to

$$\begin{aligned} \mathcal{A}^n h(x) &= h(x) + x + \frac{n\kappa}{2} + (n-1)\frac{M}{2}, \qquad n = 1, 2, \ldots, \\ u_n(x) &= x, \qquad n = 0, \ldots, N-1. \end{aligned}$$

Gdy $1-\kappa > 0$, to

$$\begin{aligned} C(x) &= \frac{\kappa M}{2} + \inf_{u\in[0,x]}\left\{(1-\kappa)\,u + \kappa(x-u) - \frac{\kappa}{2M}(x-u)^2\right\} \\ &= \frac{\kappa M}{2} + \tilde{C}(x), \end{aligned}$$

gdzie

$$\tilde{C}(x) = \begin{cases} (1-\kappa)x & \text{gdy } x < \min\{(2-\frac{1}{\kappa})2M, M\}, \\ \kappa x - \frac{\kappa}{2M}x^2 & \text{gdy } x \in [\min\{(2-\frac{1}{\kappa})2M, M\}, M], \end{cases}$$

przy czym infimum jest osiągane odpowiednio dla $u = x$ i dla $u = 0$. Stąd, gdy $\kappa < 1$, to

$$u_{N-2}(x) = \begin{cases} x & \text{gdy } x \in [0, \min\{(2 - \frac{1}{\kappa})2M, M\}], \\ 0 & \text{gdy } x \in (\min\{(2 - \frac{1}{\kappa})2M, M\}, M]. \end{cases}$$

Oczywiście dla $\kappa < 2/3$, $u_{N-2}(x) = 0$ dla wszystkich $x \in [0, M]$. Policzenie dalszych iteracji $\mathcal{A}^n h$ w przypadku $\kappa < 1$ pozostawiamy czytelnikowi.

Zadanie 18.8 Tutaj przestrzenią stanów jest $[0, \infty) \times [0, \infty)$, sterowany proces Markowa $X = (Y, Z)$. Niech P będzie operatorem przejścia dla Y. Wówczas operator przejścia P^u dla X wynosi

$$P^u v(y, z) = Pv(*, z + u)(y).$$

Zysk jest postaci

$$J_N(\pi, Y_0, Z_0) = \mathbb{E}\left\{\sum_{n=0}^{N-1} \lambda^n q(Y_n, Z_n, u_n) + \lambda^N r(Z_N, Y_N)\right\},$$

gdzie

$$q(y, z, u) = \min\{y, z\} - bu, \qquad r(y, z) = \min\{y, z\}.$$

Funkcja q przyjmuje wartości ujemne. Nie przeszkadza to jednak w zastosowaniu ogólnej teorii Bellmana bo q jest ograniczone z dołu. Dla przedziału nieskończonego dodatkowo musi być $\lambda < 1$. Operator Bellmana jest równy

$$\mathcal{A}v(y, z) = \sup_{0 \le u \le M} \{\min\{y, z\} - bu + \lambda P v(*, z + u)(y)\}.$$

Jeśli $\{Y_n\}$ jest ciągiem niezależnych zmiennych losowych o rozkładzie wykładniczym

$$\mathbb{P}(Y_n \le a) = 1 - \exp\{-a\}, \qquad a \in [0, \infty),$$

to

$$\mathcal{A}v(y, z) = \sup_{0 \le u \le M} \left\{\min\{y, z\} - bu + \lambda \int_0^\infty v(\xi, z + u) \exp\{-\xi\} d\xi\right\}.$$

Ponieważ

$$\begin{aligned} P^u \min\{y, z\} &= \int_0^\infty \min\{\xi, z + u\} \exp\{-\xi\} d\xi \\ &= \int_0^{z+u} \xi \exp\{-\xi\} d\xi + (z + u) \int_{z+u}^\infty \exp\{-\xi\} d\xi \\ &= \int_0^{z+u} \exp\{-\xi\} d\xi \\ &= 1 - \exp\{-(z + u)\}. \end{aligned}$$

Stąd

$$\begin{aligned}\mathcal{A}r(y,z) &= r(y,z) + \sup_{0\le u\le M}\{-bu+\lambda-\lambda\exp\{-(z+u)\}\}\\ &= r(y,z)+\lambda-\inf_{u\in[0,M]}\{bu+\lambda\exp\{-z-u\}\}\\ &= r(y,z)+1-b\min\{(-z-\log(b\lambda^{-1}))^+,M\}\\ &\quad -\lambda\exp\{-z-\min\{(-z-\log(b\lambda^{-1}))^+,M\}\},\end{aligned}$$

przy czym sterowanie u jest równe

$$u=\min\{(-z-\log(b\lambda^{-1}))^+,M\}.$$

Oczywiście, gdy $b\lambda^{-1}\ge 1$ to $u=0$. Stąd gdy $b\lambda^{-1}\ge 1$ to optymalnym sterowaniem na dowolnym przedziale (skończonym lub nie) jest sterowanie zerowe, to jest brak inwestycji na poprawe wydajności. Ciekawszy jest przypadek $b\lambda^{-1}<1$. Przyjmiemy, że M jest dostatecznie duże, to znaczy, że $b\lambda^{-1}\ge\exp\{-M\}$. Wtedy

$$(-z-\log(b\lambda^{-1}))^+\le -\log(b\lambda^{-1})\le M,\qquad \forall\, z\ge 0.$$

Stąd

$$u_{N-1}(y,z)=(-z-\log(b\lambda^{-1}))^+$$

oraz

$$\begin{aligned}\mathcal{A}r(y,z) &= r(y,z)+\lambda-b(-z-\log(b\lambda^{-1}))^+\\ &\quad -\lambda\exp\{-z-(-z-\log(b\lambda^{-1}))^+\}.\end{aligned}$$

Następnie

$$\mathcal{A}^2r(y,z)=r(y,z)+\lambda+\lambda^2-h_\lambda(z),$$

gdzie

$$h_\lambda(z)=-\inf_{u\in[0,M]}\{bu+\lambda\exp\{-z-u\}$$

$$+\lambda b(-z-u-\log(b\lambda^{-1}))^+ +\lambda^2\exp\{-z-u-(-z-u-\log(b\lambda^{-1}))^+\}\}$$

Zadanie 18.9 Niech $X=(X^1,\ldots,X^M)$ oznacza stan procesu, to znaczy $X^i\in\{0,1,\ldots,\}\cup\{\infty\}$ jest wiekiem i-tego bezpiecznika. Tutaj $X^i=\infty$ gdy bezpiecznik nie działa. Tak więc przestrzenią stanów jest

$$E=(\{0,1,\ldots\}\cup\{\infty\})^M.$$

Przestrzenią sterowań jest

$$U=\{0,1\}^M.$$

Sterowanie $u=(u^1,\ldots,u^M)\in U$ interpretujemy w następujący sposób: $u^i=0$ oznacza, że i-ty bezpiecznik pozostawiamy, natomiast $u^i=1$ oznacza zamiane i-tego bezpiecznika na nowy. Ponieważ bezpieczniki niedziałające

muszą być natychmiast wymienione mamy następujące zbiory dopuszczalnych sterowań
$$U(x^1,\dots,x^M)=\{u\in U:\ u^i=1 \text{ gdy } x^i=\infty\}.$$
Sterowane prawdopodobieństwa przejścia są dane wzorem

$$\mathbb{P}^u(X_{n+1}=y|X_n=x)=\prod_{i=1}^M \mathbb{P}^{u_i}\left(X_{n+1}^i=y^i|X_n^i=x^i\right),\qquad u\in U(x),$$

gdzie dla $i=1,\dots,M$, $n=0,1,\dots$ mamy

$$\mathbb{P}^1\left(X_{n+1}^i=0|X_n^i=j\right)=1,\qquad j\in\{0,1,\dots\}\cup\{\infty\}$$

oraz
$$\mathbb{P}^0\left(X_{n+1}^i=j+1|X_n^i=j\right)=\sum_{k=j+1}p_k:=q_j,\qquad j\neq\infty,$$
$$\mathbb{P}^0\left(X_{n+1}^i=\infty|X_n^i=j\right)=p_j,\qquad j\neq\infty.$$
Dla sterowania u niech
$$\varDelta(u)=\{i:\ u_i=0\}.$$
Wówczas sterowany operator przejścia jest postaci

$$P^u\psi(x)=\sum_{\varDelta\subset\varDelta(u)}P_{\varDelta,u}(x)\psi(\tau_{\varDelta,u}x),$$

gdzie $\tau_{\varDelta,u}:E\to E$ dane jest wzorem

$$(\tau_{\varDelta,u}x)_j=\begin{cases}0 & \text{gdy } j\notin\varDelta(u),\\ \infty & \text{gdy } j\in\varDelta,\\ x_j+1 & \text{gdy } j\in\varDelta(u)\setminus\varDelta\end{cases}$$

oraz
$$P_{\varDelta,u}(x)=\prod_{i\in\varDelta}p_{x_i}\prod_{i\in\varDelta(u)\setminus\varDelta}q_{x_i}.$$

Funkcjonał kosztu na przedziale skończonym $[0,N]$ jest równy

$$J_N(\pi,X_0)=\mathbb{E}\sum_{n=0}^{N-1}q(X_n,u_n)$$

gdzie, dla $u\in U(x)$,

$$\begin{aligned}q(x,u)&=a\#\{i:\ x_i\neq\delta,\ u_i=1\}+a\#\{i:\ x_i=\delta\}+b\chi_A(x)\\&=a\#\{i:\ u_i=1\}+b\chi_A(x)\\&=a\left(M-\#\varDelta(u)\right)+b\chi_A(x)\end{aligned}$$

oraz

$$A = \{x \in E : \quad x_i = \infty \text{ dla jakiegoś } i\}.$$

Stąd operator Bellmana ma postać

$$\begin{aligned}\mathcal{A}v(x) &= aM + b\chi_A(x) \\ &\quad + \inf_{u \in U(x)} \left\{ -a\#\Delta(u) + \sum_{\Delta \subset \Delta(u)} P_{\Delta,u}(x) v(\tau_{\Delta,u} x) \right\}.\end{aligned}$$

Niech $\eta(x) := \#\{i : \ x_i = \infty\}$. Wówczas

$$\mathcal{A}(0) = a\eta + b\chi_A,$$

a otymalną strategią jest

$$(u_{N-1}(x))_i = \begin{cases} 0 & \text{gdy } x_i \neq \infty, \\ 1 & \text{gdy } x_i = \infty, \end{cases}$$

czyli konieczne naprawienie tylko niedziałających bezpieczników. Zaóważmy, że

$$\sum_{\Delta \subset \Delta(u)} P_{\Delta,u}(x) = 1, \qquad x \in E, \ u \in U(x).$$

Ponadto

$$\chi_A(\tau_{\Delta,u}(x)) = 1 \quad \text{i} \quad \eta(\tau_{\Delta,u}(x)) = \#\Delta, \qquad x \in E, \ \Delta \neq \emptyset$$

oraz

$$\chi_A(\tau_{\emptyset,u} x) = 0 = \eta(\tau_{\emptyset,u}(x)), \qquad x \in E.$$

Stąd

$$\begin{aligned}&-a\#\Delta(u) + \sum_{\Delta \subset \Delta(u)} P_{\Delta,u}(x) \left(b\chi_A(\tau_{\Delta,U} x) + a\eta(\tau_{\Delta,u} x) \right) \\ &= -a\#\Delta(u) + \sum_{\Delta \subset \Delta(u), \ \Delta \neq \emptyset} P_{\Delta,u}(x)(b + a\#\Delta) \\ &= -a\#\Delta(u) + \big(1 - P_{\emptyset,u}(x)\big)\, b + a \sum_{\Delta \subset \Delta(u)} P_{\Delta,u}(x)\#\Delta.\end{aligned}$$

Strategia u_{N-2} realizuje dla każdego $x \in E$ następujące minimum

$$\begin{aligned}&\min_{u \in U(x)} \left\{ -a\#\Delta(u) + \big(1 - P_{\emptyset,u}(x)\big)\, b + a \sum_{\Delta \subset \Delta(u)} P_{\Delta,u}(x)\#\Delta \right\} \\ &= \min_{u \in U(x)} \left\{ \big(1 - P_{\emptyset,u}(x)\big)\, b - a \sum_{\Delta \subset \Delta(u)} P_{\Delta,u}(x) \left(\#\Delta(u) - \#\Delta \right) \right\} \\ &= \min_{u \in U(x)} \left\{ \big(1 - P_{\emptyset,u}(x)\big)\, b - a \sum_{\Delta \subset \Delta(u)} P_{\Delta,u}(x)\# \left(\Delta(u) \setminus \Delta \right) \right\}.\end{aligned}$$

Na nieskończonym przedziale czasowym mamy

$$J_\infty(\pi, X_0) = \mathbb{E}\sum_{n=0}^{\infty} \lambda^n q(X_n, u_n).$$

Operatorem Bellmana jest

$$\begin{aligned}\mathcal{A}_\lambda v(x) &= aM + b\chi_A(x) \\ &\quad + \inf_{u\in U(x)} \left\{ -a\#\Delta(u) + \lambda \sum_{\Delta\subset\Delta(u)} P_\Delta(x) v(\tau_\Delta x) \right\}.\end{aligned}$$

Szukamy $v_\infty = \lim \mathcal{A}_\lambda^n(0)$. Optymalna strategia stacjonarna u_∞ powinna być postaci

$$u_\infty^i(x) = \begin{cases} 1 & \text{gdy } x_i \geq \overline{x}, \\ 0 & \text{gdy } x_i < \overline{x}, \end{cases}$$

gdzie $\overline{x} < \infty$ odpowiednio dobrany poziom. Stąd

$$\Delta(u_\infty(x)) = \{i:\ x_i < \overline{x}\}.$$

Funkcjonałem kosztu na jednostkę czasu jest

$$J(\pi, X_0) = \limsup_{N\to\infty} \frac{1}{N}\mathbb{E}\sum_{n=0}^{N-1} q(X_n, u_n).$$

Rozważmy teraz szczególny przypadek dwóch bezpieczników z prawdopodobieństwami $p_1 = p_2 = 0.5$. Rozważamy przedział $[0, 3]$ czyli $N = 3$. Wtedy możemy przyjąc, że przestrze'n stanów jest równa

$$E = \{0, 1, 2, \delta\}^2,$$

przestrzeń sterowań $U = \{0, 1\}^2$. Ponieważ $r = 0$ mamy

$$\mathcal{A}r(x) = \min\{q(x, u):\ u \in U(x)\}.$$

Stąd

$$\mathcal{A}r(x) = \begin{cases} 0 & , \text{gdy } x \notin A, \\ a\#\{i:\ x_i = \delta\} + b, & \text{gdy } x \in A\} \end{cases}$$

oraz

$$u_2(x) = \begin{cases} (0,0) & \text{gdy } x \notin A, \\ (1,0) & \text{gdy } x_1 = \delta \text{ i } x_2 \neq \delta, \\ (0,1) & \text{gdy } x_1 \neq \delta \text{ i } x_2 = \delta, \\ (1,0) & \text{gdy } x_1 = \delta = x_2. \end{cases}$$

Następnie

$$\mathcal{A}^2 r(x) = \min\{q(x,u) + P^u \mathcal{A} r(x)\}.$$

Oczywiście można oganiczyć się
Zadanie 18.10 Niech Y_n będzie liczbą neutronów w chwili n przy sterownaiach $(\alpha_0, \alpha_1, \ldots, \alpha_{n-1})$. Mamy następujące prawdopodobieństwo przejścia

$$\mathbb{P}(Y_{n+1} = y | Y_n = x) = \sum_{y_1+\ldots+y_x=y} \prod_{i=1}^{x} p_{y_i}(\alpha_i).$$

Można zadać też Y w postaci rekurencyjnej

$$Y_{n+1} = \sum_{i=0}^{Y_n} \xi^i_{n+1}(\alpha),$$

gdzie $\{\xi^i_n(\alpha)\}$ ciąg iid, którego każdy element ma rozkład $p(\alpha)$.

Wobec istnienia pamięci urządzenia sterującego jako rozważamy parę

$$(Y_n, \alpha_n).$$

Przesterzenią sterowań jest $U = \{-1, 0, 1\}$. Dynamika drugiej składowej X_n jest następująca

$$\alpha_{n+1} = \alpha_n + u_n.$$

Musimy zapisać funkcjonał kosztu w postaci

$$J(\pi, X_0) = \mathbb{E} \sum_{n=0}^{\infty} q_n(X_n, u_n).$$

Wydaje się to być niemożliwe. Potrzebne jest dalsze rozszerzenie przesteni stanów. Mianowicie, definiujemy $V_{-1} = 0$, $V_0 = Y_0$,

$$V_{n+1} = \max\{V_n, Y_{n+1}\}.$$

Wtedy $(Y_n, V_n, V_{n-1}, \alpha_n)$ jest procesem Markowa i

$$\mathbb{E}\,\tau = \mathbb{E} \sum_{n=0}^{\infty} \chi_{[0,M)}(V_n)$$

oraz

$$\mathbb{P}\left(\sup_{0 \le n \le \tau} Y_n \ge \frac{3M}{2}\right) = \mathbb{E} \sum_{n=0}^{\infty} \chi_{[0,M)}(V_n) \chi_{[3M/2,\infty)}(V_{n+1}).$$

Zadanie 18.11 Naszym procesem Markowa jest (Y_n, Z_n, X_n, ξ_n). Tutaj

$$\xi_n = \begin{cases} 1 & \text{gdy w momencie } n \text{ kapitał jest ulokowany w przedsiewzięciu 1,} \\ 2 & \text{gdy w momencie } n \text{ kapitał jest ulokowany w przedsiewzięciu 1.} \end{cases}$$

Tak więc

$$X_{n+1} = \begin{cases} X_n(1+Y_n) & \text{gdy } \xi_n = 1, \\ X_n(1+Z_n) & \text{gdy } \xi_n = 2. \end{cases}$$

Przestrzenią sterowań jest $U = \{0,1\}$. Jedynie dynamika ξ zależy od sterowania i wynosi

$$\xi_{n+1} = \begin{cases} \xi_n & \text{gdy } u_n = 0, \\ f(\xi_n) & \text{gdy } u_n = 1. \end{cases}$$

Tutaj $f(1) = 2$ i $f(2) = 1$. Ponieważ proces X ma niemalejące trajektorie, funkcjonał kosztu wynosi

$$\mathbb{E}\sum_{n=0}^{\infty}\left\{\chi_{\{[0,x)\}}(X_n) + cu_n\chi_{\{[0,x)\}}(X_n)\right\}.$$

Zauważmy, że druga część odpowiada za transfery kapitału z jednego przedsięwzięcia do drugiego.

Zadanie 18.12

Literatura

1. A. Arapostathis, V.S. Borkar, E. Fernández-Gancherand, M.K. Ghoshand, and S.I. Marcus, *Discrete-time controlled Markov processes with average cost criterion: A survey*, SIAM J. Control Opt. **31** (1993), 282–344.
2. K.J. Aström, *Introduction to Stochastic Control Theory*, Academic Press, 1970.
3. R. Bellman, *Dynamic Programing,* Princeton University Press, Princeton, 1957.
4. D.P. Bertsekas, *Dynamic Programming and Stochastic Control*, Academic Press, 1976.
5. F.T. Bruss, *On an optimal selection problem of Cowan and Zabczyk,* J. Appl. Probab. **24** (1987), 918–928.
6. A. Cayley, *Problem No. 4528,* The Educational Times **27** (1874), 189–237.
7. Z. Ciesielski, J. Zabczyk, *A note on a selection problem,* Probability Theory (Z. Ciesielski, ed.), Banach Center Publication 5, Polish Scientific Publishers, 1979, 47–51.
8. R. Cowan, J. Zabczyk *An optimal selection problem associated with the Poisson process,* Theory Probab. Appl. **23** (1978), 606–613.
9. T.S. Ferguson, *Who solved the secretary problem*, Statistical Science **4** (1989), 282–296.
10. P.R. Halmos, *Measure Theory*, Van Nostrand, 1950.
11. R. Howard, *Dynamic Programing and Markov Processes*, MIT Press, Cambridge Massachusetts, 1960.
12. N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, North-Holland, Amsterdam, 1981.
13. K. Kuratowski, *Topology I*, Academic Press, New York, 1966.
14. H. Kushner, *Wprowadzenie do Teorii Sterowania Stochastycznego*, PWN, Warszawa, 1983.
15. H. Kushner, *Stochastic stability and Control*, Academic Press, New York, 1967.
16. K.P. Parthasarathy, *Probability Merasures on Metric Spaces*, Academic Press, New York, 1967.
17. P. Samuelson, *Liftime portfolio selection by the dynamic stochastic programing*, Rev. Econon Stoch. (1969), 239–246.
18. A. N. Shirjajev, *Optimal Stopping Rules*, Springer, 1978.
19. A. Sierociński, J. Zabczyk, *On a packng problem*, Bull. Pol. Acad. Sci. Math. **37** (1989), 305–313.

20. J. Zabczyk, *A selection problem associated to a renewal process*, New Trends in Systems Analysis (A. Bensoussan, J.L. Lions, eds.), Lecture Notes Control Inf. Sci. 2, Springer-Verlag, 1977, pp. 508–515.
21. J. Zabczyk, *Chance and Decision,* Quaderni, Scuola Normale Superiore, Pisa, 19996.