ACTA UNIVERSITATIS NICOLAI COPERNICI
EKONOMIA XLII – NAUKI HUMANISTYCZNO-SPOŁECZNE – ZESZYT 403 – TORUŃ 2011

Uniwersytet Mikołaja Kopernika w Toruniu
Katedra Ekonometrii i Statystyki

Sławomir Mentzen

PROCESY DECYZYJNE MARKOWA A USTALANIE KURSU W KANTORZE WYMIANY WALUT

Zarys treści. W pracy opisano teoretyczne podstawy procesów decyzyjnych Markowa (MDP), przedstawiono rynek kantorów znajdujących się na toruńskim Starym Mieście, ze szczególnym uwzględnieniem badanego kantoru, oraz opisano prosty model kantoru, który zbadano przy użyciu MDP. Znaleziono optymalne polityki ustalania kursu dla dwóch walut – euro i dolara.

Słowa kluczowe: procesy decyzyjne Markowa, kantor, optymalne decyzje.

WSTĘP

Powszechne wydaje się twierdzenie, że stosowanie zaawansowanych narzędzi optymalizacyjnych ma sens jedynie w dużych przedsiębiorstwach, które obracają wielkimi pieniędzmi. W niniejszej pracy autor postara się wykazać, że nawet bardzo małe przedsiębiorstwa mogą wykorzystać zdobycze matematyki, by proces podejmowania niektórych decyzji przestał zależeć od uczuć i emocji, przestał być subiektywny, a stał się obiektywny i pozwalał na znaczące zwiększenie zysków.

W pierwszej części artykułu opisano teoretyczne podstawy procesów decyzyjnych Markowa (MDP – Markov Decision Processes), następnie przedstawiono wykorzystany algorytm oraz zaprezentowano prosty model kantoru walutowego. Opisano też kształt rynku wymiany walut na toruńskim Starym Mieście. W ostatniej części pracy pokazano wyniki zastosowania procesów decyzyjnych Markowa do ustalenia kursów kupna i sprzedaży dwóch walut w jednym z toruńskich kantorów walutowych.

1. PODSTAWY TEORETYCZNE MODELU MDP

(X, $B(X)$ nazywany jest przestrzenią borelowską, jeśli X jest borelowskim podzbiorem zupełnej i ośrodkowej przestrzeni metrycznej i ($B(X)$ jest rodziną zbiorów borelowskich. Dla przestrzeni borelowskich (X, $B(X)$ i (Y, $B(Y)$ q jest prawdopodobieństwem warunkowym Y na X, jeśli dla $x \in X$, $q(\cdot|x)$ jest miarą probabilistyczną na $B(Y)$ i dla $G \in B(Y)$, $q(G|\cdot)$ jest mierzalną funkcją borelowską z X w [0;1]. Ścisłą definicję MDP można znaleźć między innymi u Putermana (2005):

Definicja 1.

Na model MDP składa się:
1. Przestrzeń borelowska (S, $B(S)$).
2. Przestrzeń borelowska (A, $B(A)$ i rodzina zbiorów A_s takich, że $A_s \in B(A)$ dla każdego $s \in S$. Niech $B(A_s)$ oznacza indukowaną z (A, $B(A)$ rodzinę podzbiorów borelowskich zbioru A_s, $s \in S$. Wymagamy też istnienia mierzalnej funkcji d odwzorowującej S w A tak, że $d(s) \in A_s$, dla każdego $s \in S$.
3. Rodzina $P(A_s)$ miar probabilistycznych na $B(A_s)$ dla każdego $s \in S$.
4. Funkcje rzeczywiste nagród $r_t(s, a)$ spełniające dla każdego $t \in T$ następujące warunki:
   a) $r_t(\cdot,\cdot)$ jest mierzalna względem $B(S \times A_s)$,
   b) $r_t(\cdot,\cdot)$ jest całkowalna dla wszystkich $q \in P(A_s)$, dla każdego $s \in S$.
5. Warunkowe prawdopodobieństwa $p_t(\cdot|s,a)$ spełniają dla $t \in T$ warunki:
   a) $p_t(G|\cdot,\cdot)$ jest mierzalna względem $B(S \times A_s)$ dla $G \in B(S)$,
   b) $p_t(G|s,\cdot)$ jest całkowalna dla każdego $q \in P(A_s)$, dla każdego $s \in S$ i $G \in B(S)$.

Powyżej S jest zbiorem stanów, w których może znaleźć się proces, s oznacza pojedynczy stan, T jest zbiorem etapów decyzyjnych, t jest etapem decyzyjnym, A jest zbiorem wszystkich możliwych akcji (decyzji), A_s zbiorem akcji możliwych do podjęcia w stanie s, a jest akcją (decyzją), d deterministyczną regułą decyzyjną, $r_t(s, a)$ jest nagrodą przyznaną, jeśli na etapie t w stanie s podjęta zostanie akcja a, q jest funkcją podejmowanie decyzji (deterministyczną albo losową), $p_t(s, a)$ jest prawdopodobieństwem przejścia na etapie t, przy podjęciu akcji a, ze stanu s do innego stanu.

Innymi słowy, przebieg procesu wygląda następująco. Na każdym etapie decyzyjnym proces znajduje się w jakimś stanie. Do każdego stanu przypisany jest zbiór możliwych do podjęcia akcji, zbiór nagród za wybranie poszczególnych akcji i zbiór prawdopodobieństw przejścia do innego stanu po podjęciu odpowiedniej decyzji. W momencie podjęcia decyzji przyznawana jest natychmiastowa nagroda, zależna jedynie od stanu, w którym znajduje się proces, etapu decyzyjnego i podjętej decyzji, następnie proces zgodnie z określonym prawdopodobieństwem przechodzi do innego stanu, gdzie na kolejnym etapie decyzyjnym znowu podejmowana jest akcja.

Celem rozpatrywania danego modelu MDP jest znalezienie takich reguł podejmowania decyzji, które będą maksymalizować wartość otrzymywanych nagród w czasie trwania procesu. W przypadku stosowania markowowskich deterministycznych reguł decyzyjnych, zbiorem reguł podejmowania decyzji nazywamy zbiór

$$D_t^{MD} = \{d : S \rightarrow A : d \text{ jest mierzalne i } d(s) \in A_s \text{ dla każdego } s \in S\}$$

Regułę taką nazywamy markowowską, ponieważ w każdym stanie i na każdym etapie decyzyjnym, decyzja podejmowana jest jedynie na podstawie wybranej akcji i stanu, w którym znajduje się proces. Nie ma znaczenia historia procesu, jedynie jego aktualna sytuacja (Puterman, 2005).

Definicja 2.

Polityką, planem lub strategią nazywamy zestaw reguł decyzyjnych określających podejmowane akcję w każdym momencie decyzyjnym. Politykę oznaczamy w sposób następujący: gdzie $d_t \in D_t^{MD}$ dla $t = 1, 2, \ldots, N-1$, $N \leq \infty$. Politykę nazywamy stacjonarną, jeśli $d_t = d$ dla każdego $t \in T$, gdzie $d_t \in D_t^{MD}$ jest ustaloną regułą (Puterman, 2005, s. 25).

Dla uproszczenia dalszego wywodu, dla potrzeb tej pracy przyjmujemy, że zbiory S i A są dyskretne, a horyzont czasowy jest dyskretny i nieskończony.

Rozważmy przestrzeń $\Omega = (S \times A)^{\infty} = \prod_{n-1}^{\infty} (S \times A)$. Elementy tej przestrzeni składają się z ciągu stanów i akcji: $\Omega \ni \omega = (s_1, a_1, s_2, a_2, \ldots)$. Zdefiniujmy zmienne losowe X_t i Y_t, przyjmujące wartości odpowiednio z S i A. $X_t(\omega) = s_t$, $Y_t(\omega) = a_t$, $t = 1, 2, \ldots$ X_t oznacza stan procesu w momencie t, a Y_t podjętą wtedy akcję.

W dalszej treści zmodyfikujemy założenia zawarte między innymi u Putermana (2005) w celu dostosowania ich do potrzeb i sytuacji rozważanych w niniejszym artykule.

Niech rozkład prawdopodobieństwa $P_1(\cdot)$ oznacza rozkład początkowy stanu procesu. Dla potrzeb tej pracy $P_1(s_1) = 1$ dla któregoś $s_1 \in S$.

Polityka $\pi = (d_1, d_2, \ldots)$ wprowadza prawdopodobieństwo P^{π} na $(\Omega, B(\Omega))$ następująco (Puterman, 2005, s. 23):

$$P^{\pi}\{X_1 = s\} = P_1(s).$$

$$P^{\pi}\{X_{t+1} = j \mid X_t = s,\ Y_t = a\} = p_t(j \mid s_t, a_t).$$

Prawdopodobieństwo pojedynczej ścieżki procesu ω = (s_1, a_1, s_2, a_2, … s_n) dla każdego $n \geq 1$ wynosi:

$$P^{\pi}\ \{s_1, a_1, s_2, a_2, \dots s_n) = P_1(s_1)\, p_1(s_2 \mid s_1, a_1)\, p_2\, (s_3 \mid s_2, a_2) \dots \\ p_{n-1}(s_n \mid s_{n-1}, a_{n-1}).$$

Korzystając z własności Markowa, otrzymujemy:

$$P^{\pi}\ \{a_t, s_{t+1}, a_{t+1}, s_{t+2}, \dots s_n | s_1, a_1, s_2, a_2, \dots s_t) = P^{\pi}\ \{a_t, s_{t+1}, a_{t+1}, s_{t+2}, \dots s_n | s_t).$$

Niech W oznacza rzeczywistą funkcję zmiennych losowych definiowaną na (Ω, B(Ω), P^{π}), zadaną wzorem:

$$W(s_1, a_1, s_2, a_2, \dots) = \sum_{t=1}^{\infty} r_t(s_t, a_t).$$

Definicja 3.

Wartością oczekiwaną funkcji W, przy polityce π, jest:

$$E^{\pi}W = \int_{\Omega} W(\omega) P^{\pi}\{\omega\} d\omega.$$

Od tego momentu przyjmujemy następujące założenia upraszczające model i jego notację:

1. Stacjonarność funkcji nagród i prawdopodobieństw przejścia. $r_t(s, a)$ i $p(j \mid s, a)$ nie zmieniają się przy przechodzeniu z jednego etapu decyzyjnego do kolejnego.
2. Otrzymywane nagrody są ograniczone, $r(s, a) \mid = M < \infty$ dla każdego $a \in A_s$ i $s \in S$.
3. Przestrzeń stanów S jest dyskretna i przeliczalna.
4. Polityka jest stacjonarna, $d_t = d$ dla każdego $t \in T$.

Niech $v^{\pi}(s)$ oznacza wartość oczekiwaną wszystkich nagród procesu, przy założeniu nieskończonego horyzontu czasowego, użycia polityki π oraz rozpoczęcia procesu od stanu s.

$$v^{\pi}(s) = E_s^{\pi}\left\{\sum_{t=1}^{\infty} r(X_t, Y_t)\right\}.$$

Ponieważ tak zdefiniowana wartość oczekiwana nie uwzględnia czynnika czasu, na ogół prowadzić będzie do wartości nieskończonych. Dlatego w dalszej części pracy stosowany będzie zdyskontowany MDP. Jest on znacznie bliższy rzeczywistości. Odpowiada za to zarówno inflacja, jak i preferencja czasowa sprawiająca, że wyżej ceni się teraźniejszy zysk od tego w przyszłości.

Definicja 4.

Wartością oczekiwaną procesu zdyskontowanego jest (Kadota i inni, 2006):

$$v_\lambda^\pi(s) \equiv \lim_{N\to\infty} E_s^\pi \left\{ \sum_{t=1}^{N} \lambda^{t-1} r(X_t, Y_t) \right\}$$

dla $0 \le \lambda < 1$.

Granica taka istnieje gdy $\sup_{s\in S} \sup_{a\in A_s} |r(s,a)| = M < \infty$. W takim przypadku można napisać:

$$v_\lambda^\pi(s) \equiv E_s^\pi \left\{ \sum_{t=1}^{\infty} \lambda^{t-1} r(X_t, Y_t) \right\}.$$

Definicja 5.

Polityką optymalną nazywamy politykę $v_\lambda^* \equiv \sup_\pi v_\lambda^*$.

Przy dotychczasowych założeniach wartość optymalnej polityki v_λ^* można wyznaczyć, korzystając z równań Bellmana:

$$v_\lambda(s) = \sup_{a\in A_s} \left\{ r(s,a) + \sum_{j\in S} \lambda p(j \mid s,a) v_\lambda(j) \right\}$$

dla $s \in S$ (Puterman, 2005, s. 147).

2. ALGORYTM

Poniżej znajduje się iteracyjny algorytm, służący do znajdowania optymalnej polityki decyzyjnej (Puterman, 2005, s. 161):

1. Należy wybrać v^0(s), ustalić $\varepsilon > 0$ i $n = 0$.
2. Dla każdego $s \in S$ należy wyliczyć v^{n+1}:
$$v_\lambda^{n+1}(s) = \max_{a\in A_s} \left\{ r(s,a) + \sum_{j\in S} \lambda p(j \mid s,a) v_\lambda^n(j) \right\}.$$
3. Jeśli $\| v_\lambda^{n+1} - v_\lambda^n \| < \dfrac{\varepsilon(1-\lambda)}{2\lambda}$, to należy przejść do punktu 4. W przeciwnym wypadku należy zwiększyć n o 1 i wrócić do punktu 2.
4. Dla każdego $s \in S$ wybrać
$$d_\varepsilon(s) \in \arg\max_{a\in A_s} \left\{ r(s,a) + \sum_{j\in S} \lambda p(j \mid s,a) v_\lambda^{n+1}(j) \right\}.$$
Arg max jest funkcją zwracającą wartości a, dla których argument funkcji osiąga wartości maksymalne.

3. RYNEK KANTORÓW W TORUNIU

Kantory na toruńskiej Starówce zlokalizowane są wzdłuż ulic: Chełmińska, Szeroka, Królowej Jadwigi, stanowiących najpopularniejszy trakt spacerowy w mieście. Zdecydowana większość osób udających się na Stare Miasto przechodzi ulicą Szeroką, będącą głównym deptakiem miasta, a dochodzi do niej ulicami Chełmińską bądź Królowej Jadwigi. Dlatego najlepszą lokalizację dla kantoru stanowi Szeroka, a następnie dwie pozostałe ulice. Wszystkie toruńskie kantory znajdujące się na wymienionej trasie stosują ten sam spread dla głównych walut (20 groszy dla euro i dolara, 30 groszy dla funta brytyjskiego), oraz mają te same ceny kupna i sprzedaży. Każdego dnia rano, jeśli kursy walut się w ciągu nocy zmieniły, można zaobserwować różnice sięgające nawet dwóch groszy, jednak w ciągu godziny wszystkie kantory ustalają ceny na jednym poziomie. Sporadycznie zdarza się, że któryś kantor ma inną cenę Jest to racjonalna taktyka, wszystkie te kantory mają dobrą lokalizację, dzięki czemu mogą osiągać wysokie zyski bez konkurowania pomiędzy sobą za pomocą ceny.

Kantorem, który musi wyłamać się z obowiązującej zmowy cenowej, jest ten znajdujący się na ulicy Mostowej, około 50 metrów od skrzyżowania z ulicą Szeroką. Ma dużo gorszą lokalizację, znajduje się na rzadko uczęszczanej, bocznej ulicy, otworzony został stosunkowo niedawno, dlatego jest zmuszony konkurować cenowo. Pierwszą próbą zdobycia swojego udziału w rynku było zmniejszenie proponowanego spreadu poniżej 20 groszy. Oznaczało to, że badany kantor miał najlepsze ceny zarówno kupna, jak i sprzedaży. Pozostałe kantory zareagowały jeszcze większym zmniejszeniem spreadu, co zapoczątkowało rujnującą wojnę cenową, z której zwycięsko mogły wyjść jedynie dysponujące zarówno lepszą lokalizacją, jak i dużo większym kapitałem kantory należące do konkurencji. Po paru dniach stało się jasne, że jest to zła droga i spready na lokalnym rynku wróciły do starego poziomu.

Drugą i tym razem skuteczną próbą utrzymania się na rynku było oferowanie najlepszych cen skupu albo sprzedaży przy zachowaniu spreadu na poziomie takim jak konkurencja. W praktyce oznacza to ustawienie ceny kupna waluty o grosz wyżej niż konkurencja, gdy w kasie kantoru jest dużo złotówek a mało waluty, oraz ustawienie ceny sprzedaży o grosz niżej, gdy kantor ma dużo waluty a mało złotówek. W dalszej części pracy ustawienie kursu o grosz wyżej od konkurencji nazywane będzie nastawieniem się na kupno, a o grosz niżej, nastawieniem się na sprzedaż. Dzięki tablicy z aktualnymi kursami znajdującej się na skrzyżowaniu ulic Mostowej i Szerokiej klienci są poinformowani o cenach w kantorze i jeżeli ich zamiar kupna lub sprzedaży waluty jest zbieżny z ustawieniem się kantoru na kupno lub sprzedaż, są skłonni nadłożyć drogi dla osiągnięcia korzystniejszego kursu wymiany niż w kantorach na głównych ulicach.

Trzeba zauważyć, że kursy kupna i sprzedaży w kantorach nie są ustawione symetrycznie względem średniej międzybankowej. Cena kupna dla euro i dolara jest na ogół 4 do 6 groszy niższa od średniej międzybankowej, a cena sprzedaży 14 do 16 groszy wyższa od średniej.

Przy opisanej wyżej strategii kluczowym zagadnieniem jest znalezienie takiej ilości waluty, przy której kantor powinien przestawić się z kupna na sprzedaż i odwrotnie. W przypadku gdy kantor nastawiony jest na skup waluty, traci wszystkich klientów chcących ją kupić oraz tych klientów, którzy chcą sprzedać większą ilość niż kantor jest w stanie kupić za posiadane złotówki. Sytuacja jest analogiczna w przypadku ustawienia się na sprzedaż. Zbyt późna zmiana kursu z kupna na sprzedaż powoduje zgromadzenie zbyt dużej ilości waluty, a co za tym idzie, dysponowanie zbyt małą ilością złotówek, by można było realizować większe, bardziej zyskowne transakcje kupna, co powoduje straty w wysokości nawet setek złotych dziennie.

4. MODEL MDP KANTORU WALUTOWEGO

Model zbudowany został na paru podstawowych założeniach, dostosowanych do lokalnych warunków. Pierwszym jest doskonała informacja na rynku. Wszyscy klienci posiadają wiedzę na temat cen oferowanych we wszystkich kantorach oraz o możliwości negocjacji cen przy większych wymienianych kwotach. Kolejnym założeniem jest racjonalność klientów. Sprzedadzą oni walutę w kantorze, który oferuje najwyższy kurs, a kupią tam, gdzie na tablicy widnieje najmniejsza liczba. Jeśli kantor nie jest w stanie zrealizować w całości transakcji z powodu zbyt małej ilości złotówek lub waluty, nie jest ona realizowana w części, klient odchodzi do konkurencji. Wszystkie kantory poza badanym utrzymują zbliżone kursy, jedynie badany kantor dla każdej waluty ustala najniższą lub najwyższą cenę. Kantor ustala ceny kupna i sprzedaży walut na podstawie dwóch informacji: ceny kupna i sprzedaży u konkurencji oraz ilości danej waluty w kasie.

Do badania wykorzystano zbiór wszystkich transakcji kupna lub sprzedaży z użyciem euro i dolara, o wartości powyżej 10€ i 10$, z pierwszego kwartału 2011 roku. Wszystkie transakcje zaokrąglono w górę do najbliższej liczby podzielnej przez 100, tj. 10$ do 100$, a 110$ do 200$ itd. Następnie każdej wartości transakcji przyporządkowano indeks ze zbioru $\{1, \ldots, 70\}$, równy wartości transakcji podzielonej przez 100.

Dla każdej transakcji kupna danej waluty o indeksie i wyliczono średnią marżę:

$$r_i = \frac{1}{n_i}\sum_{j=1}^{n_i} r_i^j,$$

gdzie r_i – średnia marża transakcji o indeksie i, n_i – ilość transakcji o indeksie i, r_i^j – marża j-tej transakcji o indeksie i.

Następnie dla każdej wartości transakcji wyliczono prawdopodobieństwo jej wystąpienia:

$$p_i = \frac{n_i}{n},$$

gdzie n – ilość wszystkich transakcji kupna.

W wyliczeniach koniecznie trzeba uwzględnić fakt, że kantor nie jest nastawiony przez równy czas na kupno i na sprzedaż. Dlatego potrzebne są dwie dane: t_{kupno} – czas, przez jaki kantor nastawiony był na kupno, $t_{sprzedaż}$ – czas, przez jaki kantor nastawiony był na sprzedaż. Dzięki temu można porównywać oczekiwane zyski z transakcji kupna i sprzedaży.

Kolejnym krokiem było obliczenie oczekiwanego zysku z transakcji o indeksie i według wzoru:

$$e_i = \frac{r_i \cdot i}{t_{kupno}},$$

gdzie e_i – oczekiwany zysk z transakcji o indeksie i.
Analogiczne obliczenia przeprowadzono dla transakcji sprzedaży.

Stanem procesu s jest ilość waluty w kasie kantoru podzielona przez 100, przy czym maksymalna ilość waluty wynosi 7000. System ma więc 71 stanów, od 0 do 70. Transakcje zawierają się pomiędzy 100 a 7000 i również są wielokrotnością stu. W każdym stanie należy podjąć decyzję, czy kurs należy ustawić na kupno, czy na sprzedaż. Zbiór dostępnych akcji jest więc dwuelementowy i stały dla wszystkich dostępnych stanów: $A = \{kupno, sprzedaż\}$, gdzie *kupno* oznacza decyzję nastawienia się na kupno, a *sprzedaż* na sprzedaż. W wypadku podjęcia decyzji o kupnie przyznawana jest natychmiastowa nagroda, zależna od stanu, w którym system się znajduję, o wartości:

$$R_s^{kupno} = \sum_{j=1}^{70-s} e_j \cdot p_j,$$

gdzie R_s^{kupno} – nagroda w stanie s po podjęciu decyzji o nastawieniu się na kupno waluty. W momencie podjęcia decyzji o ustawieniu kursu ma kupno wiadomo, że przyjdzie klient, który będzie chciał sprzedać walutę. Dlatego nagrodę stanowi suma oczekiwanych zysków z możliwych do zrealizowania w danym stanie transakcji, pomnożonych przez prawdopodobieństwo przyjścia klienta z daną ilością waluty. Analogicznie dla podjęcia decyzji o sprzedaży:

$$R_s^{sprzedaż} = \sum_{j=1}^{s} e_j \cdot p_j,$$

gdzie $R_S^{sprzedaż}$ – nagroda w stanie s po podjęciu decyzji o nastawieniu się na sprzedaż waluty.

Prawdopodobieństwo przejścia ze stanu i do j jest równe:

$$p(j \mid i, kupno) = \begin{cases} 0 & dla \quad j \le i \\ p_{j-i} & dla \quad j > i \end{cases},$$

oraz

$$p(j \mid i, sprzedaż) = \begin{cases} 0 & dla \quad j \ge i \\ p_{i-j} & dla \quad i > j \end{cases}.$$

Podsumowując, nieskończony proces przebiega następująco. W chwili startu w kasie znajduje się s sztuk waluty. Podejmowana jest decyzja o ustawieniu kursu na kupno lub sprzedaż. Po podjęciu decyzji przyznawana jest nagroda w wysokości R_S^{kupno} lub $R_S^{sprzedaż}$, zależna jedynie od podjętej decyzji i stanu, w którym znajduje się proces. Następnie proces przechodzi do kolejnego stanu zgodnie z prawdopodobieństwem przejścia. Celem jest maksymalizacja otrzymanych nagród, będących zyskiem kantoru. Przyszłe zyski dyskontowane są o wartość λ.

5. WYNIKI

Badanie przeprowadzono dla dwóch walut: euro i dolara. Dane pochodzą z pierwszego kwartału 2011 roku. Dane potrzebne do wyliczenia t_{kupno} oraz $t_{sprzedaż}$ pochodzą tylko z marca 2011 i są równe ilości godzin, jakie kantor nastawiony był na kupno i na sprzedaż. Brano pod uwagę tylko dwie z trzech głównych walut, ponieważ w objętym badaniem okresie kantor zawarł umowę na sprzedaż dużych ilości funtów brytyjskich, w związku z czym nastawiony był cały czas na kupno, co uniemożliwiało wiarygodne badanie. Ponieważ pełna tabela z danymi dla jednej tylko waluty miałaby ponad 70 wierszy i zajęła sporo miejsca, w artykule umieszczono tylko jej 21 pierwszych wierszy.

Dla waluty euro $t_{kupno} = 73,5$, $t_{sprzedaż} = 147,5$, $\lambda = 0,999$ $\varepsilon = 0,001$

Tabela 1. Dane kupna i sprzedaży dla euro

Kupno					Sprzedaż				
s	r	p	e	R_{kupno}	s	r	p	e	$R_{sprzedaż}$
0	0	0,00	0,00	5,90	0,00	0,00%	0,00	0,00	0,00
100	1,64%	0,52	1,17	5,90	100	2,70%	0,45	0,82	0,82
200	1,48%	0,17	0,68	5,90	200	2,56%	0,12	0,41	1,23
300	1,44%	0,08	0,50	5,90	300	2,17%	0,12	0,53	1,76
400	1,35%	0,04	0,26	5,90	400	2,39%	0,05	0,30	2,06
500	1,29%	0,04	0,38	5,90	500	2,33%	0,07	0,58	2,63
600	1,10%	0,01	0,11	5,90	600	1,38%	0,02	0,09	2,73
700	1,20%	0,01	0,11	5,90	700	2,04%	0,01	0,12	2,85
800	1,32%	0,01	0,12	5,90	800	1,39%	0,01	0,11	2,96
900	0,94%	0,01	0,08	5,90	900	1,31%	0,00	0,03	2,99
1000	1,13%	0,03	0,51	5,90	1000	1,11%	0,03	0,24	3,23
1100	0,61%	0,01	0,05	5,75	1100	1,60%	0,01	0,13	3,36
1200	1,20%	0,00	0,08	5,75	1200	1,41%	0,01	0,07	3,43
1300	1,37%	0,00	0,10	5,75	1300	1,18%	0,00	0,04	3,47
1400	0,35%	0,00	0,01	5,75	1400	0,00%	0,00	0,00	3,47
1500	1,02%	0,01	0,14	5,75	1500	0,95%	0,02	0,18	3,65
1600	1,01%	0,00	0,03	5,75	1600	0,71%	0,00	0,02	3,67
1700	0,61%	0,00	0,04	5,75	1700	1,57%	0,00	0,08	3,75
1800	0,70%	0,00	0,02	5,75	1800	0,03%	0,00	0,00	3,75
1900	0,00%	0,00	0,00	5,67	1900	0,98%	0,00	0,03	3,77
2000	0,80%	0,02	0,33	5,67	2000	1,39%	0,01	0,24	4,01

Źródło: obliczenia własne.

Dla dolara $t_{kupno} = 127,5$, $t_{sprzedaż} = 90,5$, $\lambda = 0,999$ $\varepsilon = 0,001$

Tabela 2. Dane kupna i sprzedaży dla dolara

Kupno					Sprzedaż				
s	r	p	e	R_{kupno}	s	r	p	e	$R_{sprzedaż}$
0	0,00%	0,00	0,00	4,86	0	0,00%	0,00	0,00	0,00
100	2,27%	0,53	0,94	4,51	100	3,89%	0,38	1,62	1,62
200	2,09%	0,16	0,51	4,51	200	3,78%	0,09	0,72	2,34
300	1,91%	0,07	0,33	4,51	300	3,49%	0,11	1,24	3,58
400	2,19%	0,03	0,23	4,51	400	2,99%	0,06	0,85	4,43
500	1,34%	0,06	0,32	4,51	500	3,80%	0,09	1,81	6,24
600	2,03%	0,01	0,13	4,51	600	2,58%	0,02	0,37	6,61
700	1,11%	0,01	0,04	4,51	700	3,15%	0,04	1,05	7,65
800	1,63%	0,01	0,07	4,51	800	2,42%	0,01	0,23	7,88
900	1,15%	0,01	0,05	4,51	900	4,05%	0,01	0,43	8,32
1000	1,90%	0,06	0,91	4,51	1000	2,16%	0,09	2,05	10,37
1100	0,00%	0,00	0,00	4,51	1100	0,00%	0,00	0,00	10,37
1200	0,00%	0,00	0,00	4,51	1200	0,98%	0,01	0,14	10,51
1300	1,58%	0,01	0,11	4,51	1300	0,00%	0,00	0,00	10,51
1400	0,00%	0,00	0,00	4,51	1400	0,00%	0,00	0,00	10,51
1500	0,00%	0,00	0,00	4,51	1500	1,39%	0,01	0,25	10,76
1600	0,00%	0,00	0,00	4,51	1600	0,00%	0,00	0,00	10,76
1700	0,59%	0,01	0,05	4,51	1700	0,00%	0,00	0,00	10,76
1800	0,00%	0,00	0,00	4,51	1800	0,77%	0,01	0,16	10,92
1900	0,00%	0,00	0,00	4,51	1900	0,00%	0,00	0,00	10,92
2000	1,45%	0,01	0,31	4,51	2000	2,90%	0,02	1,38	12,30

Źródło: obliczenia własne.

Po zastosowaniu opisanego w pracy algorytmu otrzymano wyniki zamieszczone w tabeli poniżej:

Tabela 3. Wartości v (s, a) dla euro i dolara

EURO			DOLAR		
s	KUPNO	SPRZEDAŻ	s	KUPNO	SPRZEDAŻ
25	**11,052**	9,738	1	**7,049**	0,000
26	**10,987**	9,738	2	**7,049**	3,772
27	**10,981**	9,738	3	**7,049**	5,016
28	**10,970**	9,738	4	**7,049**	5,869
29	**10,948**	9,937	5	7,049	**7,675**
30	**10,905**	9,937	6	7,049	**8,043**
31	**10,824**	9,937	7	7,049	**9,091**
32	**10,760**	9,954	8	7,049	**9,321**
33	**10,745**	9,954	9	7,094	**9,754**
34	**10,716**	9,954	10	7,094	**11,807**
35	**10,661**	9,954	11	7,115	**11,807**
36	**10,556**	9,979	12	7,120	**11,947**
37	**10,510**	9,979	13	7,122	**11,947**
38	**10,508**	9,979	14	7,161	**11,947**
39	**10,502**	10,067	15	7,176	**12,194**
40	**10,399**	10,067	16	7,181	**12,194**
41	**10,202**	10,067	17	7,209	**12,194**
42	**10,202**	10,067	18	7,219	**12,359**
43	**10,202**	10,067	19	7,437	**12,359**
44	**10,202**	10,129	20	7,519	**13,737**
45	**10,202**	10,129	21	7,550	**13,737**
46	**10,202**	10,129	22	7,643	**13,737**
47	**10,180**	10,129	23	7,733	**14,262**
48	**10,151**	10,129	24	7,767	**14,616**
49	10,115	**10,279**	25	7,779	**14,616**
50	10,098	**10,279**	26	7,784	**14,616**
51	9,767	**10,290**	27	7,786	**14,616**
52	9,767	**10,290**	28	7,787	**14,616**
53	9,743	**10,290**	29	7,787	**14,616**
54	9,704	**10,290**	30	7,787	**14,616**
55	9,674	**10,290**	31	7,787	**14,616**

Pogrubiono większe wartości v (s, a), wskazujące na optymalną decyzję – d (s)
Źródło: obliczenia własne.

W przypadku euro optymalną regułą decyzyjną jest ustawienie się na kupno, gdy w kasie jest nie więcej niż 4800 euro, oraz ustawienie się na sprzedaż po przekroczeniu tej kwoty. Dla dolara otrzymano następującą regułę decyzyjną: jeśli tylko ilość dolarów w kasie przekroczy 400, należy przestawić się na sprzedaż dolarów. Ta zaskakująca polityka wynika z faktu, że większość klientów kupuje małe ilości dolarów, na których kantor osiąga wysoką średnią marżę.

PODSUMOWANIE

Przeprowadzone badanie dało bardzo ciekawe i konstruktywne wyniki. Okazało się, że do wymiany euro należy używać zupełnie innej strategii niż do wymiany dolara. O ile w przypadku euro kantor powinien głównie skupywać walutę i później spróbować ją sprzedać, o tyle w przypadku dolara strategia powinna być zupełnie odmienna. Realizowanie otrzymanej strategii oznaczałoby trzymanie w kasie małych kwot, które uniemożliwiałyby większość transakcji. Dlatego kantor powinien zaopatrzyć się w banku w dużą ilość dolarów, ponosząc mały przy dużych kwotach koszt przewalutowania złotówek, a następnie cały czas sprzedawać dolary, po czym ponownie zakupić dolary w banku. W przypadku euro nie ma konieczności korzystania z usług banku.

Badanie pokazało też, że należy korzystać z narzędzi optymalizacyjnych przy każdej skali działalności, nie tylko w przypadku obrotów idących w miliony złotych. Powinno stanowić to zachętę dla mniejszych przedsiębiorców, by również zainteresowali się matematycznym wspomaganiem procesu podejmowania decyzji w ich przedsiębiorstwie.

LITERATURA

Ching W., Ng M. (2006), *Markov Chains: Models, Algorithms and Applications*, Springer, New York.

Decewicz A. (2011), *Probabilistyczne modele badań operacyjnych*, Oficyna Wydawnicza SGH w Warszawie, Warszawa.

Kadota Y., Kurano M., Yasuda M. (2006), *Discounted Markov Decision Processes*, „An International Journal Computers & Mathematics With Applications", 51.

Puterman M. (2005), *Markov Decision Processes: Discrete Stochastic, Dynamic Programming*, John Wiley and Sons, New Jersey.

Rudnicki R. (2001), *Wykłady z analizy matematycznej*, Wydawnictwo Naukowe PWN, Warszawa.

MARKOV DECISION PROCESSES AND DETERMINING THE EXCHANGE RATE

Abstract. The paper describes the theoretical foundations of Markov decision processes (MDP), presents the exchange market located in Torun's Old Town, with particular emphasis on tested exchange, and describes a simple model for an exchange, which was examined using the MDP. As a result, found the optimal policy for determining the exchange rate for two currencies – the euro and the dollar.

Key words: Markov decision processes, exchange, optimal decisions