STUDIA EKONOMICZNE

AKADEMIA EKONOMICZNA
im. Karola Adamieckiego
w Katowicach

50
ZESZYTY NAUKOWE

ZASTOSOWANIE METOD MATEMATYCZNYCH W EKONOMII I ZARZĄDZANIU

ZESZYTY NAUKOWE

AKADEMII EKONOMICZNEJ IM. KAROLA ADAMIECKIEGO

„Studia Ekonomiczne”

ZASTOSOWANIE METOD MATEMATYCZNYCH W EKONOMII I ZARZĄDZANIU

Katowice 2008

Redaktor naukowy

Jerzy Mika

Recenzent

Antoni Smoluk

Redaktor

Karolina Koluch

Korektor

Patrycja Keller

ISBN 978-83-7246-544-3

Wydawnictwo Akademii Ekonomicznej
im. Karola Adamieckiego w Katowicach
ul. 1 Maja 50, 40-287 Katowice, tel. 032 257-76-35, fax 032 257-76-43
www.ae.katowice.pl, e-mail: wydawucz@ae.katowice.pl

SPIS TREŚCI

Bartłomiej Jabłoński

ALTERNATYWNE SPOSOBY KONSTRUKCJI PORTFELI PAPIERÓW WARTOŚCIOWYCH

Wstęp

Doradcy inwestycyjni pracujący w funduszach inwestycyjnych czy też departamentach asset management starają się w taki sposób skonstruować portfele papierów wartościowych, aby osiągnąć jak najlepsze wyniki. Do podstawowych problemów, z jakimi mają do czynienia, należą:

- dobór odpowiednich walorów do portfela inwestycyjnego według analizy fundamentalnej oraz technicznej,
- określenie wielkości pozycji, czyli jakie kwoty aktywów przeznaczyć na wcześniej wybrane walory.

W przypadku wyboru odpowiednich walorów do portfela inwestycyjnego pomocna jest analiza fundamentalna, która ukazuje determinanty wyceny akcji.

Analiza fundamentalna jest złożonym procesem obejmującym kilka etapów. Najczęściej wyróżnia się:

- analizę makroekonomiczną,
- analizę sektorową,
- analizę sytuacji spółki,
- analizę finansową spółki,
- wycenę akcji [10, s. 85].

Rozpatrując problem tworzenia portfela spośród wcześniej wytypowanych akcji, przeważnie stosuje się podejście zaproponowane przez Markowitza. Idea budowania portfeli inwestycyjnych według Markowitza opiera się na wyznaczaniu portfeli efektywnych, czyli maksymalizujących oczekiwaną stopę zwrotu przy jednoczesnej minimalizacji ryzyka.

Zgodnie z pracami Markowitza, podczas konstruowania portfela papierów wartościowych największą wagę przywiązuje się do jakościowych korzyści osiąganych przez dywersyfikację inwestycji w papiery wartościowe. Model Markowitza jest oparty na metodach ilościowych. Jego podstawowe założenia są następujące:

- stopa zwrotu inwestycji należycie wyraża czerpane z niej dochody, a inwestorzy znają rozkład prawdopodobieństwa osiągnięcia danych stóp zwrotu,
- szacunki inwestorów dotyczące ryzyka są proporcjonalne do rozkładu oczekiwanych stóp zwrotu,
- inwestorzy decydują się na oparcie swoich decyzji wyłącznie na dwóch parametrach funkcji rozkładu prawdopodobieństwa, czyli na spodziewanej stopie zwrotu i prawdopodobieństwie jej osiągnięcia,
- inwestorzy skłaniają się do podejmowania minimalnego ryzyka przy danej stopie zwrotu, natomiast przy danym stopniu ryzyka wybierają projekt o największej rentowności [27, s. 75].

Model wyboru portfela akcji Markowitza, mimo że teoretycznie jest bardzo atrakcyjny, to praktycznie trudny do zastosowania. W marcu 2002 roku na Giełdzie Papierów Wartościowych w Warszawie notowano 231 spółek. Oznacza to, że chcąc skorzystać z modelu Markowitza, konieczna jest znajomość 231 stóp zwrotu, 231 odchyleń standardowych od stóp zwrotu oraz 26 565 (231*(231-1)/2) wartości współczynników korelacji. Liczby te stawiają pod znakiem zapytania możliwość zastosowania klasycznej teorii wyboru portfela akcji w sposób bezpośredni, zwłaszcza że liczba akcji na rynku będzie rosła w miarę postępowania procesu prywatyzacji w Polsce [27, s. 82].

Biorąc pod uwagę stopień skomplikowania procedur obliczeniowych, jakie trzeba wykonać, aby móc dokonać wyboru optymalnej ilości poszczególnych akcji do portfela papierów wartościowych metodą Markowitza, należy zastanowić się nad alternatywami. Metody zarządzania wielkością pozycji zaproponowane przez K. van Tharpa stanowią alternatywę w obszarze konstrukcji portfeli papierów wartościowych składających się z akcji różnych firm.

Metody te z punktu widzenia zarządzającego aktywami będą możliwe do zastosowania wobec zasobniejszych portfeli z możliwością wykorzystania dźwigni finansowej dzięki nabyciu części akcji na kredyt. Należy podkreślić, że metody te dają odpowiedź na pytanie, jaki procentowy udział poszczególnych akcji powinien składać się na cały portfel inwestycyjny, a nie jakie walory wybrać.

1. Wariancja, odchylenie standardowe oraz średni prawdziwy zakres zmiany jako miary ryzyka

Z inwestowaniem w akcje nierozerwalnie wiąże się ryzyko, które jest podstawową kategorią nowoczesnych finansów, a w szczególności problematyki inwestowania w akcje.

Niemożliwe jest jednak określenie jednoznacznego podejścia do ryzyka. Według podejścia Krzysztofa Jajugi, podstawowe rodzaje ryzyka to:

- ryzyko stopy procentowej (interest rate risk),
- ryzyko kursów walut (foreign exchange risk),
- ryzyko rynku (market risk, bull-bear market risk),
- ryzyko niedotrzymania warunków (default risk),
- ryzyko zarządzania (management risk),
- ryzyko biznesu (business risk),
- ryzyko finansowe (financial risk),
- ryzyko bankructwa (bankruptcy risk),
- ryzyko płynności (liquidity risk, marketability risk),
- ryzyko zmiany ceny (holding period risk),
- ryzyko reinwestowania (reinvestment risk),
- ryzyko wykupu na żądanie (call risk, callability risk),
- ryzyko zmienności (convertibility risk),
- ryzyko polityczne (political risk),
- ryzyko wydarzeń (event risk) [10, s. 99-101].

Rozważając ryzyko nie w odniesieniu do konkretnej instytucji, lecz ogólnie jako ryzyko finansowe, stwierdzono, iż zarządzanie ryzykiem następuje poprzez:

- poznanie ryzyka, czyli identyfikację,
- pomiar i analizę ryzyka [30, s. 161].

Powszechnie stosowaną metodą pomiaru ryzyka jest wariancja stopy zwrotu (variance of returns), nazywana również krótko wariancją. Określona jest ona za pomocą następującego wzoru:

$$V = \sum_{i=1}^{m} p_i (R_i - R)^2 \tag{1}$$

gdzie:
R – oczekiwana zmiana kursu,
R_i – i-ta możliwa do osiągnięcia wartość zmiany kursu,
p_i – prawdopodobieństwo osiągnięcia pewnej zmiany kursu,
V – wariancja stopy zwrotu [10, s. 102].

Popularnym sposobem kwantyfikacji ryzyka jest także odchylenie standardowe, określone według wzoru:

$$\sigma = \sqrt{\sum_{i=1}^{N} \left(K_i - \overline{K}\right)^2 P_i} \tag{2}$$

gdzie:
N – liczba obserwacji,
K – oczekiwany dochód,
P_i – prawdopodobieństwo i-tego dochodu,

K_i – możliwy i-ty dochód,
σ – odchylenie standardowe stopy zwrotu [7, s. 73).

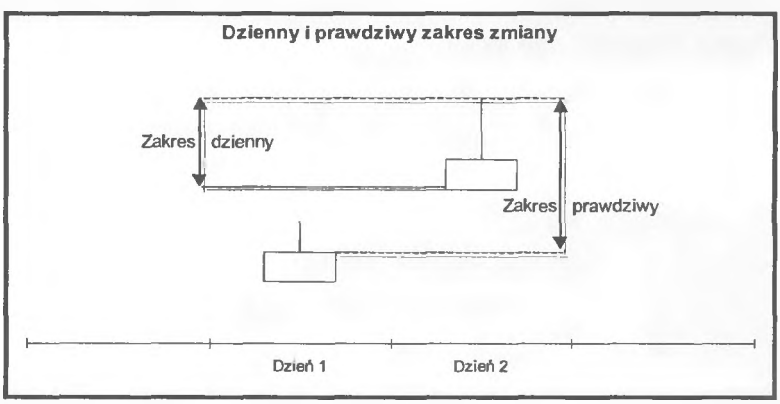
Ryzyko mierzy się na podstawie rozrzutu stóp zwrotu wokół wartości oczekiwanej. Im większe fluktuacje stóp zwrotu, tym większe ryzyko. Zróżnicowanie może być mierzone za pomocą odchylenia standardowego [7, s. 73].

Ryzyko inwestycji można także mierzyć za pomocą zmienności. Problem zmienności został omówiony przez Wildera, który opracował koncepcję prawdziwego zakresu zmiany (true range – TR). Prawdziwy zakres zmiany to najwyższa z następujących wielkości:

1. Odległość między dzisiejszym maksimum a minimum.
2. Odległość między wczorajszą ceną zamknięcia a dzisiejszym maksimum.
3. Odległość między wczorajszą ceną zamknięcia a dzisiejszym minimum.

Prawdziwy zakres zmiany sam w sobie jest po prostu liczbą. By zyskała ona sens, trzeba uśrednić wartość TR z kilku dni i stworzyć w ten sposób średni prawdziwy zakres zmiany (average true range – ATR). Wzrost tego wskaźnika oznacza wzrost zmienności [12, s. 148].

Rysunek 1 przedstawia graficzne podejście do liczenia wskaźnika ATR na przykładzie dwóch sesji. Pokazuje on sposób liczenia zakresu dziennego oraz zakresu prawdziwego.



Rys. 1. Dzienny i prawdziwy zakres zmiany

Źródło: Opracowanie własne na podstawie: [12].

Inwestycje na rynku kapitałowym są nierozerwalnie związane z ryzykiem. Mówi się, że ryzyko inwestowania w obligacje jest bardzo niskie, przy akcjach wspomina się o dużym ryzyku, zaś opcje i kontrakty terminowe są powszechnie uznawane za kwintesencję ryzyka. Takie podejście – pojawiające się, niestety, niemal w każdej pracy o inwestowaniu na rynku kapitałowym – sugeruje, jakoby

ryzyko było zawarte w samym rynku. Tymczasem należałoby powiedzieć, że ryzyko zależy od zachowania się osoby, która podejmuje jakąś działalność oraz zachowania wielu tysięcy innych inwestorów. Ryzyko nie zależy od gry, w której będziemy uczestniczyć, lecz od naszego nastawienia do niej. Można powiedzieć, że ryzyko jest funkcją chciwości i rozsądku [31, s. 85].

2. Założenia wyboru akcji do portfeli papierów wartościowych

Sposoby wyboru akcji, które mogą mieć duży potencjał wzrostowy, mogą przybierać przeróżny charakter.

Dla celów porównania modeli zarządzania wielkością pozycji opracowano sygnał nabycia akcji według reguły przecięcia średniej kroczącej przez kurs akcji, czyli proste wybicie z kanału określone przez prostą średnią kroczącą opóźnioną o jeden okres, czyli dzień.

Prostą średnią kroczącą oblicza się dodając do siebie ceny z badanego okresu, a następnie oblicza się ich średnią. Gdy pojawia się nowa wartość, wartość najstarsza wypada z wzoru. Wzór na prostą średnią kroczącą przyjmuje postać:

$$MA_t = \frac{(P_t + P_{t-1} + P_{t-2} + ... + P_{t-n})}{n} \tag{3}$$

gdzie:
MA_t – bieżąca wartość średniej kroczącej,
P_t, P_{t-1} – ceny sprzed n okresów,
n – liczba okresów użytych do obliczeń [12, s. 240].

Formuła będzie przyjmować postać:

Close>ref(Mov(c,45,s),-1)

Jest to prosta średnia krocząca liczona na podstawie notowań z 45 dni. Jako ochronę danej pozycji założono, że początkowym momentem sprzedania z zyskiem lub ze stratą w przypadku załamania się kursu akcji będzie przekroczenie poziomu 3% wartości portfela. Natomiast w przypadku ruchu w oczekiwanym przez zarządzającego kierunku będzie to linia podążającego stopu, tzw. Trailing Stop. Jest to zmienny poziom ceny, przy którym powinna nastąpić sprzedaż, jeśli cena spadającego kursu akcji osiągnie daną wartość.

Analizę oparto na notowaniach spółek wchodzących w skład indeksu WIG 20. Spośród spółek wchodzących w skład indeksu wybrano spółki z różnych branż, których cena wybiła się ponad średnią ruchomą z 45 dni przesuniętą

o 1 dzień wstecz. Zakres badanych danych obejmuje notowania od czerwca do września 2005 roku. Dla porównania wyników portfeli przedstawiono także wynik benchmarku w postaci samego indeksu spółek blue chips.

3. Wybór akcji wchodzących w skład portfeli inwestycyjnych

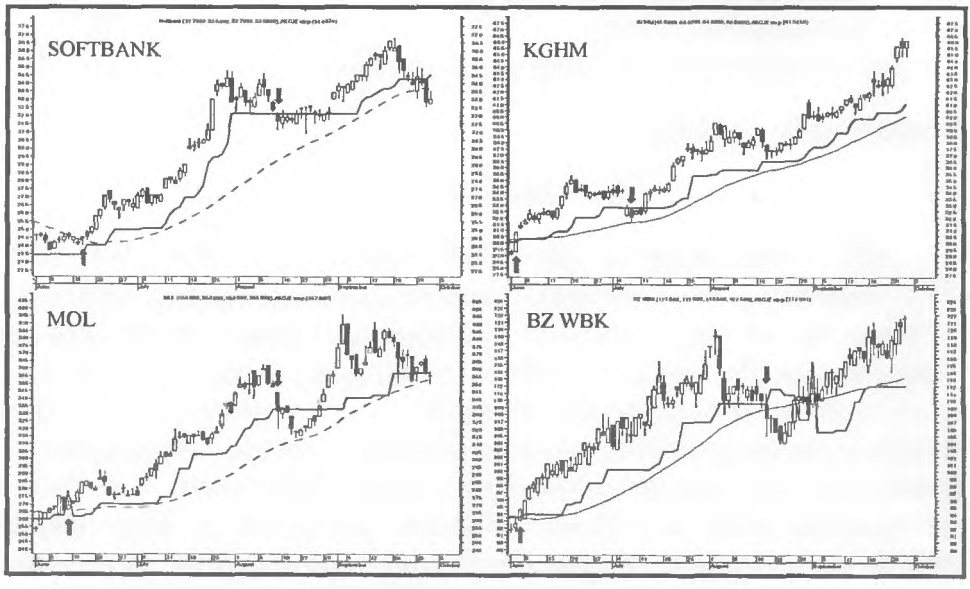
Spośród spółek wchodzących w skład indeksu WIG 20 wybrano te, które według reguły średniej kroczącej dały sygnał nabycia. Tabela 1 prezentuje dane spółek wraz z indeksem WIG 20.

Tabela 1

Wykaz badanych instrumentów

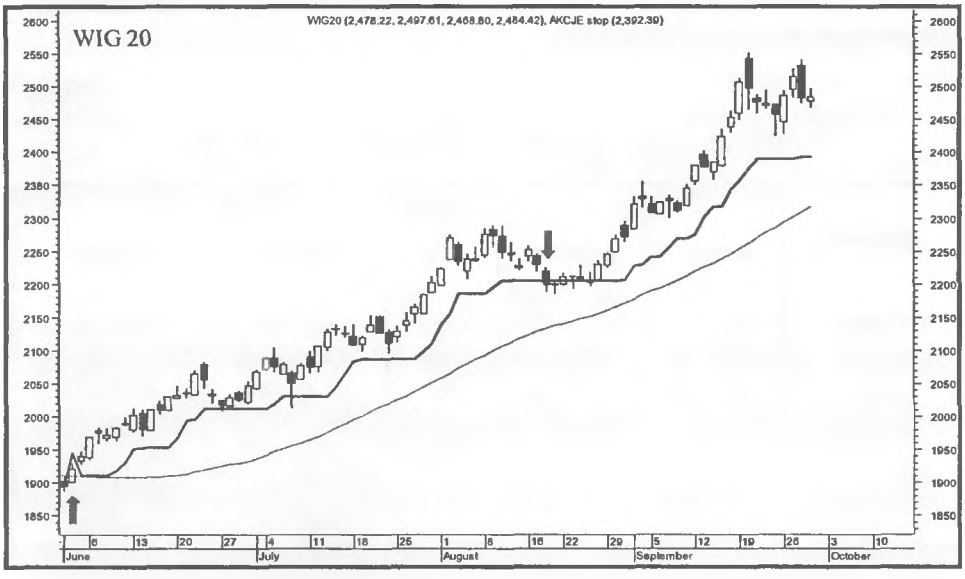
Instrument	Cena (zł)		Okres	Zysk	
	zakupu	sprzedaży		(zł)	(%)
BZWBK	94,50 zł	107,50 zł	czerwiec - sierpień	13,00 zł	13,8%
KGHM	30,80 zł	32,00 zł	czerwiec - lipiec	1,20 zł	3,9%
MOL	269,00 zł	331,50 zł	czerwiec - sierpień	62,50 zł	23,2%
SOFTBANK	24,50 zł	31,60 zł	czerwiec - sierpień	7,10 zł	29,0%
WIG 20	1 921,00 zł	2 198,00 zł	czerwiec - sierpień	277,00 zł	14,4%

Rysunek 2 przedstawia akcjogram spółek wraz z ukazanymi dniami zakupu oraz sprzedaży akcji. Linia gładka (przerywana) prezentuje średnią kroczącą, a szarpana (ciągła) poziom, przy którym należy pozbyć się akcji z portfela.



Rys. 2. Wykresy wytypowanych papierów wartościowych

Rysunek 3 przedstawia, jak zachowywał się kurs indeksu WIG 20 będący benchmarkiem dla modelowych portfeli papierów wartościowych



Rys. 3. Wykres indeksu giełdowego WIG 20

4. Alternatywne modele budowania portfeli papierów wartościowych

Omawiane modele konstrukcji portfeli papierów wartościowych K. van Tharpa scharakteryzowano w następujący sposób:
- model pierwszy – jednostki o równej wartości,
- model drugi – strategia ryzyka procentowego,
- model trzeci – model procentowej zmienności.

Dla celów porównawczych analizy wielkość środków własnych przeznaczonych do konstrukcji modeli portfelów będzie stała i równa 500 tys. zł. W obliczeniach nie uwzględniono prowizji od nabycia i zbycia pakietów akcji.

Model pierwszy – jednostki o równej wartości

Budowa modelu opartego na jednostkach o równej wartości polega na podziale całego portfela na równe części i przeznaczeniu tych części aktywów na zakup poszczególnych walorów. Wartość aktywów przeznaczoną na dany walor dzieli się przez cenę przypuszczalnego nabycia i w wyniku otrzymuje się liczbę walorów do zakupu.

Tabela 2 prezentuje wyliczenia liczby akcji, jakie należy nabyć stosując model konstrukcji portfela na podstawie jednostek o równej wartości. Zaprezentowano również wartość zakupu oraz sprzedaży pakietów wraz z ukazaniem zyskowności wszystkich transakcji.

Tabela 2

Wynik pierwszego modelu portfela papierów wartościowych

Instrument	Cena		Liczba	Wartość	
	zakupu	sprzedaży	akcji	zakupu	sprzedaży
BZWBK	94,50 zł	107,50 zł	1322	124 929,00 zł	142 115,00 zł
KGHM	30,80 zł	32,00 zł	4058	124 986,40 zł	129 856,00 zł
MOL	269,00 zł	331,50 zł	464	125 085,00 zł	154 147,50 zł
SOFTBANK	24,50 zł	31,60 zł	5102	124 999,00 zł	161 223,20 zł
			Sumy:	499 999,40 zł	587 341,70 zł
				Wynik (zł):	**87 342,30 zł**
				Wynik (%):	**17,47%**

Model drugi – strategia ryzyka procentowego

W momencie kiedy zarządzający portfelem inwestycyjnym zamierza zakupić walory, powinien określić poziom, przy którym w razie niepowodzenia akcje zostaną sprzedane, aby chronić kapitał przed dalszą jego aprecjacją. Jest to ryzyko pozycji, czyli strata, jaką ponosimy w trakcie najgorszego scenariusza.

W przypadku modelu strategii ryzyka procentowego należy określić ryzyko, które w tym modelu powinno zależeć od szerokości stopów stosowanych w celu ochrony kapitału. Zatem na samym początku należy założyć, ile możemy zaryzykować w każdej transakcji w odniesieniu do całego portfela, po czym ustalić dla każdego waloru, ile możemy zaryzykować na spadku ceny. Dla celów analizy założono dopuszczalną stratę w wysokości 2% wartości portfela.

Na każdej transakcji ryzykujemy 2% wartości początkowej portfela, czyli:

$$2\% * 500\,000 \text{ zł} = 10\,000 \text{ zł}$$

Każdorazowo liczy się ilość akcji, jaką można nabyć w ramach strategii tworzenia portfela papierów wartościowych według reguły:

$$\text{Ryzyko na każdej transakcji/stop początkowy} = \text{liczba akcji} \qquad (4)$$

W przypadku zastosowania wielkości możliwej początkowej straty (Trailing Stop) w modelu strategii ryzyka procentowego często wartość zakupu przewyższa początkowy kapitał, co wymaga wykorzystania kredytu na zakup akcji. W przypadku kiedy doradca inwestycyjny trafnie wytypuje grupy akcji do portfela inwestycyjnego, wykorzystanie kredytu znacznie powiększy stopę zwrotu osiągniętą z inwestycji. W modelu założono koszt kredytu na poziomie 12% w skali roku. Z uwagi na długość inwestycji całego portfela (czerwiec-sierpień) obliczono koszt kredytu dla 3 miesięcy. Sytuację przedstawia tabela 3.

Tabela 3

Wynik drugiego modelu portfela papierów wartościowych

Instrument	Cena		Stop	Liczba	Wartość	
	zakupu	sprzedaży		akcji	zakupu	sprzedaży
BZWBK	94,50 zł	107,50 zł	3,50 zł	2857	269 986,50 zł	307 127,50 zł
KGHM	30,80 zł	32,00 zł	0,99 zł	10101	311 110,80 zł	323 232,00 zł
MOL	269,00 zł	331,50 zł	7,00 zł	1428	384 132,00 zł	473 382,00 zł
SOFTBANK	24,50 zł	31,60 zł	1,10 zł	9090	222 705,00 zł	287 244,00 zł
				Sumy:	1 187 934,30 zł	1 390 985,50 zł
					Kwota kredytu (zł):	687 934,30 zł
					Roczny koszt kredytu (%):	12%
					Kwartalny koszt kredytu (zł):	20 638,03 zł
					Wynik (zł):	**182 413,17 zł**
					Wynik (%):	**36,48%**

W przypadku wykorzystania kredytu na zakup akcji zysk na całym portfelu znacznie wzrasta. Jednocześnie ze wzrostem zaangażowania w portfelu kapitału obcego rośnie ryzyko całego portfela. W przypadku nieodpowiedniego zabezpieczenia portfela przed stratami, dodatkowe wykorzystanie kredytu na zakup akcji spowoduje znaczne zmniejszenie jego wartości.

Model trzeci – model procentowej zmienności

Trzeci model tworzenia portfela papierów wartościowych opiera się na wskaźniku zwanym zmiennością. Zmienność to przeciętna wielkość dziennej zmiany instrumentu bazowego w określonym czasie. Jest to bezpośrednia miara zmienności cen, z jaką możemy mieć do czynienia w przypadku danej pozycji. Jeśli porówna się zmienność wszystkich utrzymywanych pozycji, przedstawiając jako procent całego kapitału, można zrównać wahania wartości każdego elemen-

tu portfela. W większości przypadków zmienność to różnica między najwyższą a najniższą ceną dnia.

Na każdej transakcji ryzykujemy 2% wartości początkowej portfela, czyli:

$$2\% * 500\,000 \text{ zł} = 10\,000 \text{ zł}$$

Każdorazowo liczy się ilość akcji, jaką można nabyć w ramach strategii tworzenia portfela papierów wartościowych według reguły:

$$\text{Ryzyko na każdej transakcji/zmienność (ATR)} = \text{liczba akcji} \quad (5)$$

Do obliczenia średniego prawdziwego zakresu zmiany ryzyka akcji wybrano 3-dniową średnią ATR (3). Przy wyborze kierowano się liczbą dni notowań w ciągu tygodnia. Dodatkowo starano się zachować w miarę krótki czas uśrednienia ze względu na małą wagę historycznych wartości wskaźnika sprzed n dni oraz dużą wagę ostatnich dni notowań. Wyniki inwestycji osiągnięte według zasad trzeciego modelu przedstawia tabela 4.

Tabela 4

Wynik trzeciego modelu portfela papierów wartościowych

Instrument	Cena		ATR	Liczba	Wartość	
	zakupu	sprzedaży	(3 dni)	akcji	zakupu	sprzedaży
BZWBK	94,50 zł	107,50 zł	1,75 zł	5714	539 973,00 zł	614 255,00 zł
KGHM	30,80 zł	32,00 zł	0,83 zł	12048	371 078,40 zł	385 536,00 zł
MOL	269,00 zł	331,50 zł	6,20 zł	1612	433 628,00 zł	534 378,00 zł
SOFTBANK	24,50 zł	31,60 zł	0,53 zł	18867	462 241,50 zł	596 197,20 zł
				Sumy:	1 806 920,90 zł	2 130 366,20 zł
				Kwota kredytu (zł):		1 306 920,90 zł
				Roczny koszt kredytu (%):		12%
				Kwartalny koszt kredytu (zł):		39 207,63 zł
				Wynik (zł):		**284 237,67 zł**
				Wynik (%):		**56,85%**

5. Porównanie modeli

W każdym modelu obowiązują rożne metody określania ilości akcji wchodzących w skład portfela, co jednocześnie przekłada się na różną wartość, jaką należy zainwestować w celu budowy portfela według jednej z trzech metod analizy portfelowej. Tabela 5 ukazuje porównanie omawianych modeli konstrukcji portfeli papierów wartościowych.

Tabela 5

Podsumowanie wyników portfeli

Portfele	Zysk (zł)	Zysk (%)
Jednostki o równej wartości	87 342,30 zł	17,47%
Strategia ryzyka procentowego	182 413,17 zł	36,48%
Model procentowej zmienności	284 237,67 zł	56,85%

Model tworzenia portfela papierów wartościowych na podstawie jednostek o równej wartości jest najprostszym sposobem konstrukcji portfela inwestycyjnego. Dwa ostatnie modele są dopasowane do większego portfela, w którym są środki przeznaczone także na dywersyfikację ryzyka poprzez inwestycję w papiery opatrzone minimalnym ryzykiem. Wtedy można część środków przeznaczyć na zakup większej ilości akcji, rezygnując po części z inwestycji w obligacje. Takie krótkotrwałe zachwianie struktury procentowej portfela opłaca się, albowiem gdyby zaciągnąć kredyt pod zakup akcji, inwestycja nie wykazywałaby już tak dużej rentowności.

Indeks WIG 20 osiągnął w badanym okresie stopę zwrotu równą 14,4%. Porównanie wyniku indeksu WIG 20 z badanymi portfelami, szczególnie drugim i trzecim modelem, ukazuje, jak wielkie znaczenie w tworzeniu portfeli papierów wartościowych ma odpowiedni dobór ilości poszczególnych pakietów akcji. Dla wytrawnego inwestora najodpowiedniejszy wydaje się być trzeci model, który – jak wynika z symulacji – łączy wysoką stopę zwrotu z wyższym ryzykiem.

Tabela 6 zawiera podsumowanie wszystkich modeli tworzenia portfeli papierów wartościowych.

Tabela 6

Wynik trzeciego modelu portfela papierów wartościowych

Model	Zalety	Wady
Model równych jednostek	– Każda jednostka zyskuje taką samą wagę w portfelu	– Drobny inwestor może powiększyć pozycję dopiero po długim czasie – Ryzyko dla poszczególnych jednostek może być równe – Nie wszystkie aktywa dają się równo dopasować do równej jednostki
Ryzyko jako procent kapitału	– Umożliwia stopniowy wzrost niezależnie od wielkości rachunku – Równoważy ryzyko w obrębie portfela, odnosząc je do określonej ryzykowanej sumy	– Należy zrezygnować z niektórych transakcji, ponieważ są zbyt ryzykowne – może to być wadą – Procentowy próg ryzyka nie musi być identyczny z faktycznym ryzykiem
Model procentowej zmienności	– Umożliwia stopniowy wzrost niezależnie od wielkości rachunku – Równoważy ryzyko w obrębie portfela, odnosząc je do zmienności – Może służyć do równoważenia transakcji w systemie stosującym bliskie linie stop	– Należy zrezygnować z niektórych transakcji, ponieważ są zbyt ryzykowne – Dzienna zmienność to nie to samo, co faktyczne ryzyko

Źródło: Opracowanie własne na podstawie: [28, s. 236].

Wnioski

Modele ukazują ważny aspekt analizy portfelowej odnoszący się do dywersyfikacji ryzyka, poprzez nabycie bezpieczniejszych od akcji walorów – obligacji. Występuje alternatywa w stosunku do kredytu na zakup akcji. Jest to możliwe w momencie, kiedy początkowo zakłada się dywersyfikację ryzyka poprzez przekazanie np. 40% wartości portfela na inwestycje obarczone mniejszym ryzykiem, czyli wspomniane obligacje. Jednak w momencie rezygnacji z przekazania kwoty na nabycie obligacji, większą część gotówki można przeznaczyć na zakup akcji, co przełoży się na niższe koszty pozyskania dodatkowego – obcego kapitału. Należy mieć na uwadze, że spowoduje to wzrost ryzyka całego portfela. Praktyka wskazuje, że w krótkich okresach podobne sytuacje są wykorzystywane przez zarządzających, szczególnie w przypadku funduszy zrównoważonych o podwyższonym poziomie ryzyka.

Analiza portfelowa – obok sztuki wyboru akcji o dużym potencjale wzrostowym – jest najważniejszą umiejętnością profesjonalnego doradcy inwestycyjnego. Jednak czas, jaki należy poświęcić na zbudowanie portfela metodą Markowitza oraz założenia, jakie są wymagane przy budowie takiego portfela powodują, że praktycznie nie jest to satysfakcjonujący sposób budowy zdywersyfikowanego portfela akcji. Zakładając częste zmiany grup akcji, konstrukcja portfela papierów wartościowych zgodnie z założeniami Markowitza jest bezpodstawna, gdyż model taki należałoby bardzo często aktualizować.

Należy brać pod uwagę inne sposoby określane także mianem analizy portfelowej, które pomogą określić procentowe zaangażowanie w poszczególne walory. Nie tylko są one mniej czasochłonne, lecz także można przypuszczać, iż dużo efektywniejsze. Przedstawione alternatywy udowadniają, że można konstruować portfele w inny sposób, aniżeli za pomocą tradycyjnych metod, dający także godziwe zyski pod warunkiem oczywiście, iż w skład tak skonstruowanego portfela będą wchodzić akcje o potencjale wzrostowym.

Literatura

1. Biegański M.: Hedging i nowoczesne usługi finansowe. Wydawnictwo Akademii Ekonomicznej w Poznaniu, Poznań 2001.
2. Bodie Z., Kane A., Markus A.J.: Investment. Irwin, Homewood 1989.
3. Dębski, W.: Rynek finansowy i jego mechanizmy. Wydawnictwo Naukowe PWN, Warszawa 2001.
4. Foster G.: Financial Statement Analysis. Englewood Cliffs 1978.

5. Francis J.C.: Inwestycje. Analiza i zarządzanie. WIG PRESS, Warszawa 2000.
6. Graham B.: Mądry inwestor. Wydawnictwo Profesjonalnej Szkoły Biznesu, Kraków 1999.
7. Groppelli A.A.: Wstęp do finansów. WIG PRESS, Warszawa 1999.
8. Haugen R.A.: Nowa nauka o finansach. WIG PRESS, Warszawa 1999.
9. Hull J.: Kontrakty terminowe i opcje. Wprowadzenie. WIG PRESS, Warszawa 1999.
10. Jajuga K., Jajuga T.: Inwestycje. Wydawnictwo Naukowe PWN, Warszawa 2002.
11. Krutsinger J.: Systemy transakcyjne. Sekrety mistrzów. WIG PRESS, Warszawa 2000.
12. LeBeau C.: Komputerowa analiza rynków terminowych. WIG PRESS, Warszawa 1999.
13. Lee T.: Ekonomia dla inwestorów giełdowych. WIG PRESS, Warszawa 2000.
14. Mantegna R.N.: Ekonofizyka. Wprowadzenie. Wydawnictwo Naukowe PWN, Warszawa 2001.
15. MetaStock – podręcznik użytkownika. Equis International, Salt Lake City 1995.
16. Murphy J.J.: Analiza Techniczna. WIG PRESS, Warszawa 1999.
17. Murphy J.J.: Międzyrynkowa Analiza Techniczna. WIG PRESS, Warszawa 1999.
18. Niederliński A.: Regułowe systemy ekspertowe. Wydawnictwo Komputerowe Jacka Skalimierskiego, Gliwice 2000.
19. Niederman D.: Wizjonerzy, sceptycy, łowcy okazji. WIG PRESS, Warszawa 2000.
20. Pieczyński A.: Reprezentacja wiedzy w diagnostycznym systemie ekspertowym. Lubelskie Towarzystwo Naukowe, Zielona Góra 2003.
21. Pring M.J.: Podstawy analizy technicznej. WIG PRESS, Warszawa 1998.
22. Schwager J.D.: Analiza techniczna rynków terminowych. WIG PRESS, Warszawa 2002.
23. Sobczyk M.: Matematyka finansowa. Podstawy teoretyczne, przykłady, zadania. Placet, Warszawa 2001.
24. Sobczyk M.: Statystyka. Wydawnictwo Naukowe PWN, Warszawa 2001.
25. Stridsman T.: Trading Systems That Work: Building and Evaluating Effective Trading Systems. McGraw-Hill Trade 2000.

26. Tadion J.M.W.: Rozszyfrować rynek. Prognozowanie, inwestowanie, wskaźniki, dane i statystyka. WIG PRESS, Warszawa 1999.
27. Tarczyński W.: Fundamentalny portfel papierów wartościowych. PWE, Warszawa 2002.
28. Tharp K. van: Giełda, wolność i pieniądze. WIG PRESS, Warszawa 2000.
29. Tharp K. van: Barton D.R., Sjuggerund S.: Bezpieczne strategie inwestycyjne. Oficyna Ekonomiczna, Kraków 2006.
30. Wąsowski W.: Ekonomia i finanse banku komercyjnego. Zarządzanie i Finanse, Warszawa 2002.
31. Zalewski G: Kontrakty terminowe w praktyce. WIG PRESS, Warszawa 2001.

ALTERNATIVE WAYS OF STRUCTURING THE SECURITIES PORTFOLIO

Summary

In the article we presented the portfolio analysis issues that show the alternative methods of the securities portfolio investment structures. By analysing the investment process, in particular the problem of the capital venture risk, we paid attention to the risk estimators. The very important problem mentioned in the article is the connection of the risk together with the size of the entry. This is the determination of the adequate volume of the given share groups included in the investment portfolio.

The article also presents the comparison of the securities portfolio creation models, concentrating in particular on both the equity and borrowed capitals. We have also included examples of the borrowed capital influence on the final outcome, which resulted from the usage of models based on the percentage and variable risks.

The article includes the analysis of the possible results gained depending on the preferred approach to the analysis of the securities portfolio creation and the assets allocated for the investment.

The study includes a theoretical description of risk estimation issues through variance, standard deviation and the true average range of changes. Thanks to the description of the portfolio analysis issues, examples presentation and description we have solved the problem of the securities portfolio creation by means of models other than that of Markowitz.

Jan Kałuski

PROGRAMOWANIE WIELOETAPOWE A *n*-OSOBOWE GRY Z KOALICJAMI

Wstęp

Pojęcie *n*-osobowej gry z koalicjami wprowadzili Neumann i Morgenstern [8]. We wczesnych latach 50. ubiegłego wieku Gelles użył pojęcia rdzenia (jądra) gry jako narzędzia do studiowania stabilności koalicji w *n*-osobowych grach. Następnie Shapley i Shubik rozwinęli to pojęcie, proponując koncepcję rozwiązania tych gier. Idea *n*-osobowych gier z koalicjami w sytuacjach bez wypłat ubocznych (transferu wypłat) należy do Shapley'a i Shubika [12] oraz Luce'a i Reiffa [11]. Idea modelowania systemów sterowania jako hierarchicznych dynamicznych gier koalicyjnych była zapoczątkowana przez Novikovą [9] oraz Krutova i Novikovą [4]. Koncepcja modelowania dyskretnych procesów przemysłowych jako wieloetapowych *n*-osobowych gier koalicyjnych należy do Kałuskiego [3].

Niniejsza praca jest poświęcona metodzie programowania wieloetapowego (MPW) w zastosowaniu do rozwiązywania *n*-osobowych gier koalicyjnych.

W związku z powyższym w punkcie 1 pracy szczegółowo przedstawiono metodę MPW, zaś w punkcie 2 na potrzeby niniejszej pracy opisano wieloetapową *n*-osobową grę koalicyjną. Powiązania pomiędzy metodą MPW i wieloetapową *n*-osobową grą koalicyjną zostały przedstawione w punkcie 3. W punkcie 4 pracy pokazano przykładowe rozwiązanie wieloetapowego problemu decyzyjnego modelowanego za pomocą wieloetapowych *n*-osobowych gier z koalicjami, stosując algorytm szeregowania metody MPW.

1. Metoda programowania wieloetapowego

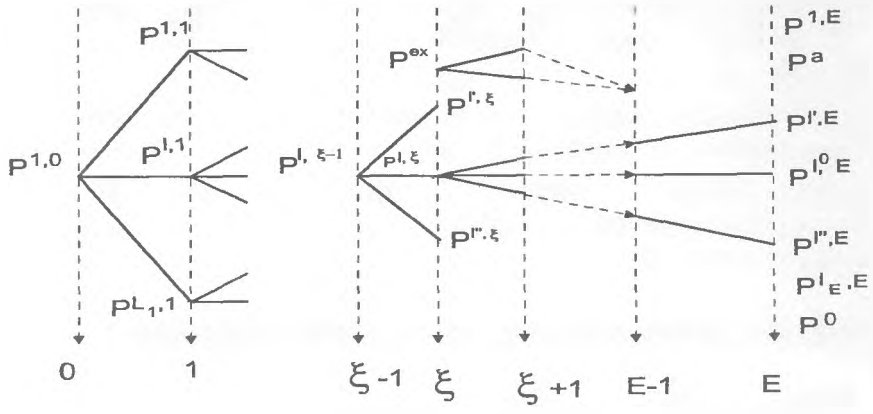
Koncepcja metody programowania wieloetapowego (MPW) pochodzi od Mareckiego [6; 7]. Metoda ta opiera się na idei programowania dynamicznego oraz metodzie podziału i ograniczeń. Istnieją 4 algorytmy tej metody: binarny,

sekwencyjny, alokacji oraz algorytm szeregowania. W wieloetapowym procesie decyzyjnym rozróżnia się: stany, decyzje oraz funkcje transformacji stanów. Dla systemu stan obrazuje sytuację w wybranej chwili. W procesach deterministycznych na podstawie danego stanu i podjętej decyzji w tym stanie za pomocą funkcji transformacji można otrzymać kolejny stan. W wyniku podejmowania decyzji w kolejnych etapach przechodzi się od stanu początkowego do stanu końcowego. W ten sposób otrzymuje się ciąg stanów, który tworzy trajektorię. Natomiast ciąg podjętych decyzji będzie strategią w danym postępowaniu decyzyjnym. Z każdego stanu można, rzecz jasna, wygenerować wiązkę trajektorii. W problemach kombinatorycznych, takich jak np. harmonogramowanie, wiązka trajektorii wychodząca ze stanu początkowego przedstawia drzewo decyzyjne. Węzłami tak powstałego drzewa są stany, zaś łukami odpowiednio decyzje powstałe w wyniku transformacji stanów. Dla wyznaczenia optymalnej trajektorii, tzn. optymalnego stanu końcowego, są generowane wszystkie możliwe trajektorie. W celu znalezienia optymalnego rozwiązania i skrócenia czasu obliczeń trajektorie nieperspektywiczne są eliminowane. Algorytmy programowania wieloetapowego są różnicowane ze względu na sposób eliminacji trajektorii nieperspektywicznych.

Algorytmy programowania wieloetapowego opierają się na odpowiednim zdefiniowaniu stanu, wartości stanu, procedury generowania stanów oraz reguł eliminacji stanów nieperspektywicznych. Na podstawie pracy Mareckiego [6] przedstawiono podstawowe definicje i procedury metody programowania wieloetapowego.

1.1. Definicja stanu

Rozpatrzono wieloetapowy proces decyzyjny, który ukazano na rys. 1.



Rys. 1. Ilustracja ogólnego algorytmu metody programowania wieloetapowego

Źródło: [6].

W tym procesie wyróżniono etapy decyzyjne $0,1,\ldots,\xi-1,\xi,\xi+1,\ldots E$. Każdy etap ξ z E-etapów ($\xi = 1,2,\ldots,E$) ma L_ξ stanów $P^{l,\xi}$. Stany są numerowane w przedziale jednego etapu ($l = 1,2,\ldots, L_\xi$). Stan procesu decyzyjnego jest macierzą o wymiarze nxd. Każdemu obiektowi ϖ_i, i=1,2,...,n jest przyporządkowany i-ty wiersz macierzy stanu. Liczba kolumn d zależy od struktury systemu. W i-tym wierszu macierzy stanu $P^{l,\xi}$ jest zapisana informacja o numerze agregatu A_m, w którym podejmuje się decyzje o obiekcie ϖ_i, sposobie realizacji tej decyzji, tj. wykonaniu operacji. Zapisywane są tu również chwile rozpoczęcia i zakończenia operacji.

1.2. Klasyfikacja stanów

W wieloetapowym procesie decyzyjnym (rys. 1) wyróżnia się następujące stany:

1. Początkowy stan $P^{1,0}$. Stan ten interpretuje początkowe warunki decyzyjnego procesu przed rozpoczęciem harmonogramowania. Stan $P^{1,0}$ jest macierzą zerową, jeżeli wszystkie obiekty, co do których podejmujemy decyzje, są przed systemem, w którym wykonuje się pewne operacje nad tymi obiektami. W przeciwnym przypadku niektóre elementy macierzy $P^{1,0}$ są dodatnie.
2. Aktywny stan $P^{l,\xi}$, tzw. stan perspektywiczny, pozwala na generację innych stanów.
3. Wybrany stan $P^{\lambda,\xi-1}$ jest stanem aktywnym, który został wybrany dla generowania następnych stanów.
4. Wygenerowanym stanem P jest stan, który został otrzymany ze stanu $P^{\lambda,\xi-1}$. Stan ten jest testowany ze względu na jego perspektywiczność, tzn. możliwość generacji następnych stanów. Nieperspektywiczny stan jest eliminowany z dalszego procesu decyzyjnego.
5. Wyczerpany stan P^{ex} jest stanem wygenerowanym P, z którego nie można otrzymać żadnego stanu końcowego.
6. Końcowy stan $P^{l,E}$ jest najlepszym stanem otrzymanym w skończonym czasie (np. w skończonym czasie obliczeń na komputerze).
7. Lokalnie optymalny stan $P^{l_0,\xi}$ jest najlepszym stanem otrzymanym ze stanu $P^{l,\xi}$.
8. Globalnie optymalny stan P^0 jest najlepszym stanem końcowym.

Podczas generowania stanów można otrzymać identyczne stany o tych samych współrzędnych. Przy tym różne strategie mogą prowadzić do różnych lub identycznych stanów.

Załóżmy, że S jest lokalnie optymalnym harmonogramem otrzymanym na podstawie wygenerowanej trajektorii. Mówimy, że dwa stany $P^{l,\xi}$ i $P^{\lambda,\xi}$ są alternatywne, jeżeli harmonogram $S^{l^0,\xi}$ otrzymany ze stanu $P^{l,\xi}$ może być zrealizowany.

1.3. Wartość stanu

Każdy końcowy stan $P^{l,E}$ wprost określa dopuszczalny harmonogram $S^{l,E}$. Stąd każdy stan $P^{l,\xi}$ dla $\xi<E$ określa harmonogram $S^{l,\xi}$. Do estymacji harmonogramu S są określone kryteria optymalizacji Q_K, $k = 1,\ldots,L$. W sposób analogiczny do estymacji stanu $P^{l,\xi}$ określa się wartość stanu $V^{l,\xi}$. Wartość stanu w tym przypadku jest zapisywana w postaci następującego wektora:

$$V^{l,\xi} = \left[\frac{1}{c} T^{l,\xi}\right] \tag{1}$$

gdzie:

c – cykl, tzn. systematyczny czas między chwilami podejmowania decyzji o wykonaniu operacji na obiekcie,

$[\bullet]$ – część całkowita.

Współrzędne tego wektora są określone na podstawie funkcji wartości $V_K(P^{l,\xi})$. Funkcje wartości odpowiadają kryterium Q_K. Znając stan $P^{l,\xi}$, możliwe jest otrzymanie jego wartości $V^{l,\xi}$. Używając wartości $V_K(P^{l,\xi})$, wydłużamy jednak czas podjęcia decyzji (wykonania obliczeń). W celu uniknięcia takiej sytuacji podczas generowania stanu P ze stanu $P^{\lambda,\xi-1}$, równocześnie obliczmy jego wartość. W tym celu korzystamy z formuł rekurencyjnych, które dla addytywnych funkcji kryterialnych obliczamy ze wzoru:

$$v_k = v_k^{\lambda,\xi-1} + \Delta v_k \tag{2}$$

z warunkiem początkowym $V_K^{1,0}$. Analogicznie można użyć formuł rekurencyjnych dla kryteriów minimalizujących i maksymalizujących. Formuła rekurencyjna (2) zmniejsza czas obliczeń (czas podjęcia decyzji w algorytmie). Wymaga to jednak zapamiętania nie tylko stanu $P^{l,\xi}$, ale również wartości tego stanu $V^{l,\xi}$.

Jest oczywiste, że dla problemu jednokryterialnego wartość stanu jest skalarem. W tym przypadku globalny optymalny stan jest wyznaczany z warunku:

$$\left(\min_{1\le l\le E} V^{l,E} = V^{l^0,E}\right) \Rightarrow \left(P^{l^0,E} = P^0\right) \tag{3}$$

Dla problemów wielokryterialnych na podstawie wartości stanu jest możliwe wyznaczenie zbioru Pareto-optymalnych stanów. Stan $P^{\lambda,E}$ dominuje nad stanem $P^{l,\xi}$, gdy spełniony jest następujący warunek:

$$\underset{1\le l\le l}{\forall}\ \underset{j\ne k}{\exists}\left(\min_{1\le l\le L_E} v_k^{\lambda,E} = v_k^{l^0,E}\right)\wedge\left(v_j^{\lambda,E} = v_j^{l,E}\right) \tag{4}$$

Zbiór stanów Pareto-optymalnych tworzą stany niezdominowane. Koncepcja dominacji stanów, zdefiniowana przez warunek (4), może być rozszerzona na stany alternatywne $P^{\lambda,E}$ i $P^{\lambda,E}$ dla etapów $\xi<E$.

Dla wyznaczenia stanu polioptymalnego jest stosowana metoda dialogowa lub sam problem wielokryterialny jest redukowany do jednokryterialnego problemu za pomocą funkcji użyteczności.

Załóżmy, że mamy do czynienia z hierarchią kryteriów Q_K i Q_L. Niech tolerancje wskaźników wynoszą q_K. Wówczas stan polioptymalny może być otrzymany z następujących zależności:

$$\underset{1\le l\le L_E}{\forall}\ \underset{1<j<l}{\exists}\ \underset{k<j}{\forall}\left(v_l^{\lambda,E} = v_l^{l,E} + q_l\right)\vee\left(\left|v_k^{\lambda,E} - v_k^{l,E}\right| \le q_k\right)\wedge\left(v_j^{\lambda,E} \le v_j^{l,E} + q_j\right)\Rightarrow\left(P^{\lambda,E} = P^0\right) \tag{5}$$

Jeżeli $q_K = 0$, można wyznaczyć tylko jeden stan polioptymalny z zależności (5). Ponadto warunek (5) pozwala określić, który stan z dwóch stanów jest lepszy, w przypadku gdy kolejno generujemy stany końcowe.

W metodzie programowania wieloetapowego stany polioptymalne są wyznaczane po jednokrotnym wygenerowaniu trajektorii.

1.4. Generowanie stanów

Celem generowania stanów jest znalezienie kompletnej wiązki trajektorii (rys. 1). Każda trajektoria ma początek w danym stanie startowym $P^{1,0}$, z którego są otrzymywane stany $P^{l,1}$, $l=1,\ldots L_1$. Ogólnie z wybranego stanu $P^{\lambda,\xi-1}$ są generowane stany etapu ξ. W ten sposób otrzymuje się stany końcowe $P^{l,E}$, $l=1,\ldots,L_E$.

Lista stanów aktywnych, reguły wyboru, reguły podziału oraz procedury generowania stanów mają podstawowe znaczenie dla generowania stanów. Do generowania stanów są wykorzystywane stany aktywne $P^{l,\xi}$ umieszczone na odpowiedniej liście. Lista stanów aktywnych jest uporządkowanym zbiorem stanów. Uporządkowanie to polega na wyborze z listy k_ξ, stanów ξ-etapu, $\xi = 0,1,\ldots E-1$, oraz odpowiednim ponumerowaniu stanów znajdujących się na tych listach. Sposób numerowania stanów aktywnych ma wpływ na efektywność

algorytmu. Generowanie stanów polega na wyborze pewnego stanu $P^{\lambda,\xi-1}$, używając go jako stanu bazowego.

Po wygenerowaniu wszystkich stanów ze stanu $P^{\lambda,\xi-1}$ stan ten jest usuwany z listy $k_{\xi-1}$. Wygenerowane stany są wprowadzane na listę k_{ξ}. Ze stanu $P^{\lambda,E-1}$ są generowane stany końcowe. Stany te nie są aktywne. Generowanie stanów jest przerywane, jeżeli lista stanów aktywnych jest pusta.

Reguły wyboru stanu aktywnego są używane do znajdowania stanu $P^{\lambda,\xi-1}$. Ze stanu tego będą generowane kolejne stany. Przykładami klasycznych reguł wyboru są znane z teorii masowej obsługi reguły FIFO, LIFO oraz LLB. Szczegółowy opis reguł wyboru można znaleźć w pracy [6].

Stosowane reguły podziału wyznaczają z kolei wiązki (pęczki) trajektorii generowanych z wybranego stanu $P^{\lambda,\xi-1}$. Z tego stanu są generowane wszystkie jego bezpośrednie następne stany (następniki) $P^{\lambda,\xi}$, jeżeli podział jest zupełny. Wówczas stan $P^{\lambda,\xi-1}$ przestaje być aktywny. Gdy dokonamy podziału częściowego, ze stanu $P^{\lambda,\xi-1}$ również otrzymamy tylko część bezpośrednich następników. Stąd stan $P^{\lambda,\xi-1}$ nadal pozostaje aktywny. Reguły częściowe pozwalają na uwzględnienie ograniczeń wielkości list stanów aktywnych. W przypadku szczególnym tylko jeden stan $P^{\lambda,\xi}$ jest generowany. Pozwala to na zapamiętanie w trakcie obliczeń tylko jednego stanu aktywnego. Zauważmy, że stosowanie reguł częściowego podziału wymaga zapamiętywania dodatkowych informacji określających stan, które nie są jednak potrzebne przy podziale całkowitym.

Procedury generowania stanów tworzą zdanie logiczne, tzn. przy spełnieniu określonego warunku logicznego można wygenerować stan. Precyzuje się przy tym warunki wyznaczania elementów macierzy nowego stanu. Wyróżniamy jednokrokowe i wielokrokowe procedury generowania stanów. W jednokrokowej procedurze ze stanu $P^{\lambda,\xi-1}$ otrzymujemy stan $P^{\lambda,\xi}$. Procedura wielokrokowa ze stanu $P^{\lambda,\xi-1}$ pozwala otrzymać stan $P^{\lambda,k}$, $k>\xi$.

1.5. Reguły eliminowania stanów

Dla zwiększenia efektywności algorytmów programowania wieloetapowego stany nieperspektywiczne są usuwane. Stan jest nieperspektywiczny, jeżeli nie pozwala na wygenerowanie rozwiązania optymalnego. W ten sposób już na wczesnych etapach generowania stanów, wiązki trajektorii nieperspektywicznych są eliminowane. W przeciwnym przypadku stan P jest umieszczany na liście stanów aktywnych.

Do usuwania stanów nieperspektywicznych wykorzystuje się reguły wyczerpywania dominacji oraz sondowania. Reguła wyczerpywania jest warun-

kiem logicznym, jaki musi spełnić stan P. Warunek ten wynika z ograniczeń procesu.

Twierdzenie 1

Jeżeli w stanie P nie można podjąć jakiejś decyzji (np. wykonania operacji) bez naruszenia ograniczeń procesu, to wyczerpane zostały możliwości generowania rozwiązań dopuszczalnych (stanów końcowych).

Skrót dowodu jest zamieszczony w pracy [6].

Reguła dominacji pozwala na usunięcie jednego z dwóch stanów aktywnych $P^{l_1,\xi}$, $P^{l_2,\xi}$, dla którego odpowiedni lokalnie optymalny stan jest gorszy. Jeżeli wspomniane stany spełniają warunek (4), to stan $P^{l_1,\xi}$ dominuje nad stanem $P^{l_2,\xi}$. Podczas obliczeń na ξ-etapie stany $P^{\lambda,E}$ i $P^{l^0,E}$ są nieznane. Wówczas decyzja o stanie zdominowanym jest podejmowana na podstawie twierdzenia 2.

Twierdzenie 2

Jeżeli stan $P^{l,\xi}$ jest alternatywny ze stanem $P^{\lambda,\xi}$ oraz $v_i^{\lambda,\xi} \le v_i^{l,\xi}$, $i = 1,...,I$, to stan $P^{\lambda,\xi}$ dominuje nad stanem $P^{l,\xi}$.

Ideę dowodu tego twierdzenia również można znaleźć w [6].

Ostatnimi regułami używanymi w programowaniu wieloetapowym są reguły sondowania. Reguła sondowania eliminuje stan $P^{l,\xi}$, z którego lokalnie optymalny stan jest gorszy od aktualnie najlepszego stanu. W tym celu dla stanu $P^{\lambda_\xi^0,E}$ wyznacza się dolne ograniczenie $d_i^{l^0,E}$. Jeżeli znany jest stan aktualnie najlepszy P^a, wówczas stan $P^{l,\xi}$ jest nieperspektywiczny, gdy:

$$\underset{1\le i\le I}{\forall} v_i^a \le d_i^{l^0,E} \tag{6}$$

Oprócz tego, jeżeli ze stanu $P^{l,\xi}$ jest wyznaczony pewien stan końcowy, spełniający ograniczenie:

$$\underset{1\le i\le I}{\forall} v_i^{k,E} \le d_i^{l^0,E} \tag{7}$$

wówczas ten stan jest lokalnie optymalny. Stąd generowanie wiązki trajektorii ze stanu $P^{l_1,\xi}$ jest niepotrzebne, gdyż stan $P^{\lambda_\xi^0,E}$ dominuje nad pozostałymi końcowymi stanami $P^{l_\xi^0,E}$.

W zakończeniu tego punktu, dla przykładu, przedstawiono algorytm szeregowania metody programowania wieloetapowego. W algorytmie tym są dane współrzędne stanu dla chwili zakończenia wykonywania operacji. Zakładamy, że:

$$P_i^{l,\varsigma} = \begin{cases} t_i, gdy operacja\ \varpi_i \text{ jest wykonywana} \\ 0, \text{ w przypadku przeciwnym} \end{cases} \tag{8}$$

oraz że przyrost wartości stanu wynosi:

$$\Delta P_j = \begin{cases} t_i \text{ dla } j = i \\ 0 \text{ w przeciwnym przypadku} \end{cases} \tag{9}$$

Ze stanu $P^0 \Rightarrow t_i \Rightarrow P_i^0$, tzn. optymalny harmonogram wykonania operacji. Chwile t_i otrzymujemy z zależności:

$$t_i = \begin{cases} T^{\lambda,\xi-1} + \vartheta_i, \text{ gdy } \left[\frac{1}{c}T^{\lambda,\xi-1}\right] = \left[\frac{1}{c}(T^{\lambda,\xi-1} + \vartheta_i)\right], \\ c\left[\frac{1}{c}T^{\lambda,\xi-1}\right] + \vartheta_i \quad \text{w przypadku przeciwnym} \end{cases} \tag{10}$$

gdzie $T^{l,\xi}$ jest chwilą zakończenia wszystkich operacji ze zbioru operacji $\Omega^{l,\xi}$ (w stanie $P^{l,\xi}$). Stąd:

$$T^{l,\xi} = \max_{\omega_i \in \Omega^{l,\xi}} t_i \tag{11}$$

Dla $P^{1,0}, T^{1,0} = 0$. Natomiast wartość stanu $V^{l,\xi}$ jest skalarem $v^{l,\xi}$ podanym wzorem:

$$v^{l,\xi} = \left[\frac{1}{c}T^{l,\xi}\right] \tag{12}$$

2. Wieloetapowa *n*-osobowa gra z koalicjami

W pracach Fudenberga i Tirole'a [1] oraz Osborne'a i Rubinsteina [10] jest omawiana klasa gier, tzw. „wieloetapowe gry z obserwowaną akcją". W tych grach:

1. Wszyscy gracze znają akcje (ruchy) na wszystkich poprzednich etapach $0,1,2,\ldots,E-1,E,E+1$, gdy wybierają akcję na etapie ξ.
2. Wszyscy gracze wykonują swoje akcje jednocześnie na każdym etapie ξ.

Pierwszym etapem w tych grach jest etap „0" dla uproszczenia notacji w przypadku, gdy etapy są interpretowane jako cykle (okresy). Mówi się, że gracze wykonują swoje akcje równocześnie na etapach ξ, jeżeli każdy gracz wybiera swoja akcję bez wiedzy na tym etapie o posunięciach pozostałych graczy. Przeciwieństwem do „jednoczesnego" wykonywania akcji przez graczy jest naprzemienne wykonywanie akcji. Dla przykładu gra Stackelberga ma 2 etapy, a gracze naprzemiennie wykonują swoje akcje.

Należy zauważyć, że często w naturalny sposób pojęcie „etapu" gry jest utożsamiane z czasem. Nie jest to jednak konieczne dla istnienia etapu, tzn. etap nie musi być utożsamiany z czasem.

Tak więc w naszym przypadku zakładamy, że wieloetapowe *n*-osobowe gry z koalicjami są podobne do „wieloetapowych gier z obserwowaną akcją". Na pierwszym etapie tej gry (etap 0) wszyscy gracze $i \in J, J = \{1,2,\ldots,n\}$ równocześnie wybierają akcję ze zbioru $\alpha_\xi, \xi = 0$ gdzie α_ξ jest określony podzbiorem zbioru J (jest to zbiór złożony ze struktur koalicyjnych na danym etapie). Na końcu każdego etapu wszyscy gracze mają możność zaobserwowania profilu wykonywanych akcji. Ogólnie akcje i-tego gracza osiągalne do realizacji na l-tym etapie mogą zależeć od akcji na poprzednich etapach.

Jako przykład omawianej gry rozpatrzono problem Balansowania Linii Montażowej (BLM) jako wieloetapową *n*-osobową grę z koalicjami. Pokazano, że problem BLM, gdzie występuje decydent centralny z ustaloną liczbą asemblerów (montażystów), jest systemem hierarchicznym, a w grze, w której centralny decydent podejmuje decyzje co do alokacji montażystów do określonych miejsc (stanowisk) linii montażowej, jest wieloetapowym problemem decyzyjnym.

Wieloetapowe hierarchiczne *n*-osobowe gry z koalicjami były zapoczątkowane przez Novikovą [9]. W swojej pierwszej pracy autorka przeanalizowała strukturę informacyjną dwuosobowej gry, a w dalszych otrzymane wyniki rozszerzyła na przypadek *n*-osobowej gry z wybranymi strukturami hierarchicznymi oraz *n*-osobowej gry z ustaloną sekwencją ruchów graczy. Rozwinięta (ekstensywna) forma zapisu gry zawiera następujące informacje:

1. Zbiór graczy.
2. Kolejność ruchów graczy.
3. Wypłaty dla graczy zależne od ruchów graczy.
4. Jakie plany mają gracze wykonując dany ruch?
5. Co każdy gracz wie, kiedy dokonuje wyboru?
6. Rozkład prawdopodobieństwa zdarzeń zewnętrznych.

W niniejszej pracy aspekty probabilistyczne w rozpatrywanej grze nie będą omawiane. Dla dalszej formalizacji rozpatrywanego problemu przytoczono pewne wyniki, które były otrzymane w pracy [3], dotyczące modelu BLM.

Teoriogrowy model funkcjonowania systemu linii montażowej z zadanymi uczestnikami gry i jednym centrum zarządzania określono w następujący sposób:

1. Zadany jest zbiór elementów systemu $J_0 = \{0,1,2,...n\}$ nazywanych dalej graczami, $|J_0| = N = n+1$. Gracza z numerem "0" oznaczamy G_0 i nazywamy decydentem. Gracz ten nie bierze bezpośredniego udziału w grze. W jego jednak interesie gra jest prowadzona. Gracze z numerami od 1 do n tworzą zbiór $J = \{1,2,3,...n\}$ graczy G_i, $i = 1,2,...,n$, którzy bezpośrednio biorą udział w grze przy linii montażowej. Każdy gracz G_i charakteryzuje się czasem wykonania ϑ_i, $i = 1,\ldots,n$, określonego zadania. Decydent ustala kolejność wykonywania poszczególnych zadań. Zadany jest cykl linii montażowej c.
2. Gra przebiega w taki sposób, że w kolejnych następujących po sobie cyklach, których liczba z góry nie jest znana, ale jest skończona, bierze udział określony podzbiór zbioru J graczy. Oznaczmy liczbę cykli potrzebnych do wykonania całej pracy przez L_ξ, $\xi = 0,1\ldots$. Liczba ta jest również liczbą etapów dynamicznej n-osobowej gry. Jeżeli nie będzie to wprowadzało nieporozumień, to niekiedy zamiast oznaczenia L_ξ będzie stosowane oznaczenie L.
3. Gracz G_i, $i = 1,..., n$ wykonuje swoje zadanie na agregacie znajdującym się na m-tym stanowisku pracy, $m = 1,...,M$. Liczba M stanowisk pracy nie jest również z góry znana.
4. Każdy gracz ma określony zbiór $V^i_{\xi,\gamma}$ strategii $v^i_{\xi,\gamma}$ w grze w zależności od numeru etapu oraz ustalonej kolejności w grze zadanej macierzą $G = [\gamma_{k,i}]$, $k,i = 1,\ldots,n$. Dla gracza G_0 będziemy pisali $v^0_\xi \in V^0_\xi$, a jego strategie nie zależą od $\gamma \in G$. Elementy macierzy G są liczbami binarnymi:

$$\gamma_{k,i} = \begin{cases} 1, \text{ jeżeli operacja } \omega_k \text{ jest bezpośrednim} \\ \quad \text{poprzednikiem operacji } \omega_i \\ 0, \text{ w przypadku przeciwnym} \end{cases} \tag{13}$$

Wieloetapowy proces przydziału zadań do stanowisk na linii montażowej może być opisany przez następujący układ równań:

$$X_{\xi+1} = f_\xi(X_\xi; v_\xi^0, v_{\xi,\gamma}^1, \ldots, v_{\xi,\gamma}^n), \quad X_\xi \in E^r \quad \xi = 0,1,\ldots \tag{14}$$

$$X_0 = A, \tag{15}$$

$$v_\xi^0 \in V_\xi^0, \quad v_{\xi,\gamma}^i \in V_{\xi,\gamma}^i, \quad i \in J, \gamma \in G, \quad \xi = 0,1,\ldots \tag{16}$$

gdzie X_ξ jest r-wymiarowym wektorem stanu systemu w chwili ξ (na etapie ξ), a $r = 0,1,2,\ldots$ jest liczbą różnych struktur koalicyjnych, którą można utworzyć na danym etapie gry. Stan $X_0 = A$ odzwierciedla stan gry na zerowym etapie gry. Jest to również stan zaawansowania prac na linii montażowej w końcowym etapie poprzedniej gry. W szczególnym przypadku $X_0 = 0$. Cele C^i graczy G_i, $i \in J_0$ są opisywane funkcjami skalarnymi stanu końcowego linii montażowej:

$$C^i(v^0, v_\gamma^1, \ldots, v_\gamma^n) = g_\gamma^i(X_{L_\xi+1}), \quad i \in J_0, \gamma \in G \tag{17}$$

gdzie $v_\gamma^i = \{v_{0,\gamma}^i, \ldots, v_{\xi,\gamma}^i\}$ to realizowane przez i-tego gracza zadanie, zaś $X = \{X_0, \ldots, X_{L_\xi+1}\}$ – odpowiadająca temu zadaniu trajektoria. Cele graczy są wyrażone ich wypłatami $g_\gamma^i(\cdot)$. Funkcja wypłat gracza G_i jest funkcjonałem w postaci:

$$g_\gamma^i = g_\gamma^i(\Psi_L(v^0, v_\gamma^1 \ldots, v_\gamma^n)), \quad i \in \{0\} \cup J; \quad J = \{1,2,\ldots,n\}, \quad \gamma \in G \tag{18}$$

określonym na zbiorze kartezjańskim $V^0 \times \cdots \times V^n$, $V^i = \prod_{\xi=0}^{L} V_{\xi,\gamma}^i, i \in J$ za pomocą odwzorowania:

$$\Psi_L \div \prod_{i \in J_0} \prod_{\xi=0}^{L} V_{\xi,\gamma}^i \to \mathbf{X}_{L+1} \tag{19}$$

gdzie $\mathbf{X}_{L+1}$ to zbiór wszystkich możliwych końcowych punktów trajektorii dynamicznego systemu. Analogicznie do (19) zdefiniowano odwzorowanie:

$$\Psi_\tau \div \prod_{i \in J_0} \prod_{\xi=0}^{\tau} V_{\xi,\gamma}^i \to \mathbf{X}_{\tau+1}, \quad \tau = 0,1,\ldots,L-1 \tag{20}$$

oraz:

$$\Psi = (\Psi_0, \ldots, \Psi_L) \tag{21}$$

gdzie $\mathbf{X}_{\tau+1}$ to zbiór wszystkich możliwych końcowych punktów $X_{\tau+1}$ kawałków $(X_0, \ldots, X_{\tau+1})$ trajektorii $X = \Psi(v)$ generowanych przez układ (14)-(16).

Założono dalej, że każdy z graczy stara się zwiększyć swoją wypłatę, używając dostępnych dla siebie strategii $v^{j}_{\xi,\gamma}$. Założono również, że gracze mają dokładną informację o parametrach linii montażowej.

Centrum uwzględnia możliwość tworzenia się koalicji poszczególnych graczy, tzn. w pierwotnej grze Γ pewne zespoły graczy (zbiory) $K_\xi \in J$, na etapie ξ, $0 \le \xi \le L$ mogą łączyć się ze sobą w koalicje $K_\xi \in \alpha_\xi$, gdzie α_ξ to ustalona klasa podzbiorów zbioru J (zbiór struktur koalicyjnych na danym etapie). Gracz G0 nie należy do żadnej koalicji. Koalicja K_ξ ma określony zbiór fizycznie dostępnych wyborów $V^{K}_{\xi,\gamma} = \prod_{i \in K_\xi} V^{i}_{\xi,\gamma}$, z którego jest wybierana $v^{K}_{\xi,\gamma} = (v^{K}_{\xi,\gamma} / i \in K_\xi)$ na zasadzie dążenia do maksymalizacji funkcji wypłat dla koalicji. Na danym etapie podzbiory koalicji są tworzone w taki sposób, że utworzone dwie różne koalicje powinny się różnić co najmniej jednym graczem lub kolejnością występowania graczy. W ten sposób:

$$K_\xi = \left\{ \left\{ K^{i}_{\xi,\gamma} / (\gamma_{k,i} = 1), k,i \in J \right\} / \underset{k,i}{\forall} \in J, K^{i}_{\xi,\gamma} \subseteq J, \gamma \in G \right\} \qquad (22)$$

gdzie $K^{i}_{\xi,\gamma}$ to koalicja, która w chwili ξ składa się z graczy G_i spełniających warunek kolejnościowy. Zbiór wszystkich możliwych podzbiorów α_ξ złożony ze struktur koalicyjnych na etapie ξ oznaczmy przez:

$$\Im_\xi = \left\{ K_\xi \in \alpha_\xi / \left\{ K^{i}_{\xi,\gamma} \subseteq J \right\}, \gamma_{k,i} \in G; k,i \in J \right\} \text{ oraz } \Im = \prod_{\xi=0}^{L} \Im_\xi$$

Wynikiem $(v;\aleph) \in V \times \Im$ dynamicznej gry z koalicjami nazwiemy ciąg $(v;\aleph) = \{v_0;\aleph_0\},\ldots,(v_L;\aleph_L\}$, gdzie $\aleph = \aleph_0,\ldots,\aleph_l)$ jest programem struktur koalicyjnych. Każdemu wynikowi $(v;\aleph)$ przyporządkowujemy wektor wypłat graczy G_i, $i \in J$. Oznaczmy wobec tego przez $g^{i,\aleph}_{\gamma}(\Psi(v))$ wypłatę gracza G_i w sytuacji, gdy $\Psi_L(v) \in \mathbf{X}_{L+1}$.

Podsumowując, hierarchiczną dynamiczną grę z koalicjami można zdefiniować w postaci uporządkowanej czwórki w następujący sposób:

$$\Gamma_\alpha = \left\langle \alpha_\xi, \left\{ \left\{ V^{K_\xi}_{\xi,\gamma} \right\}_{K_\xi \in \alpha_\xi, \gamma \in G} \right\}_{\xi=0}^{L}, \left\{ g_G{}^{\aleph} \right\}_{\aleph \in \alpha_\xi}, \left\{ \Im_\xi \right\}_{\xi=0}^{L} \right\rangle \qquad (23)$$

W tej grze nie ma możliwości dodatkowych wypłat dla graczy.

3. Powiązanie między metodą programowania wieloetapowego a wieloetapową *n*-osobową grą koalicyjną

W poprzednich punktach pracy sformułowano metodę programowania wieloetapowego jako wieloetapowego procesu podejmowania decyzji oraz wieloetapową *n*-osobową grę z koalicjami. Łatwo zauważyć, że taka gra jest także wieloetapowym procesem decyzyjnym, w którym na każdym etapie gry gracze podejmują decyzję o przynależności do określonej koalicji. Można sformułować wobec tego następujące dwa wnioski:

Wniosek 1

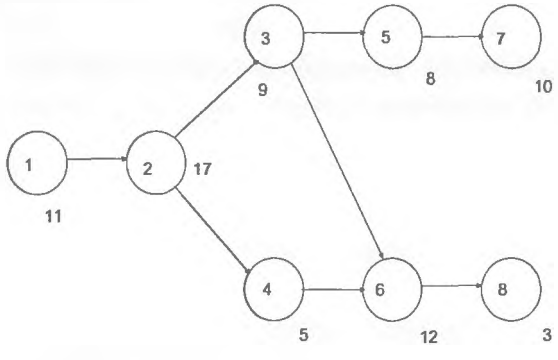
Każda wieloetapowa *n*-osobowa gra z koalicjami generuje wieloetapowy proces decyzyjny, wobec tego grę taką można rozwiązać metodą programowania wieloetapowego.

Wniosek 2

Metoda programowania wieloetapowego generuje wieloetapowy proces decyzyjny, który jest wieloetapową *n*-osobową grą koalicyjną.

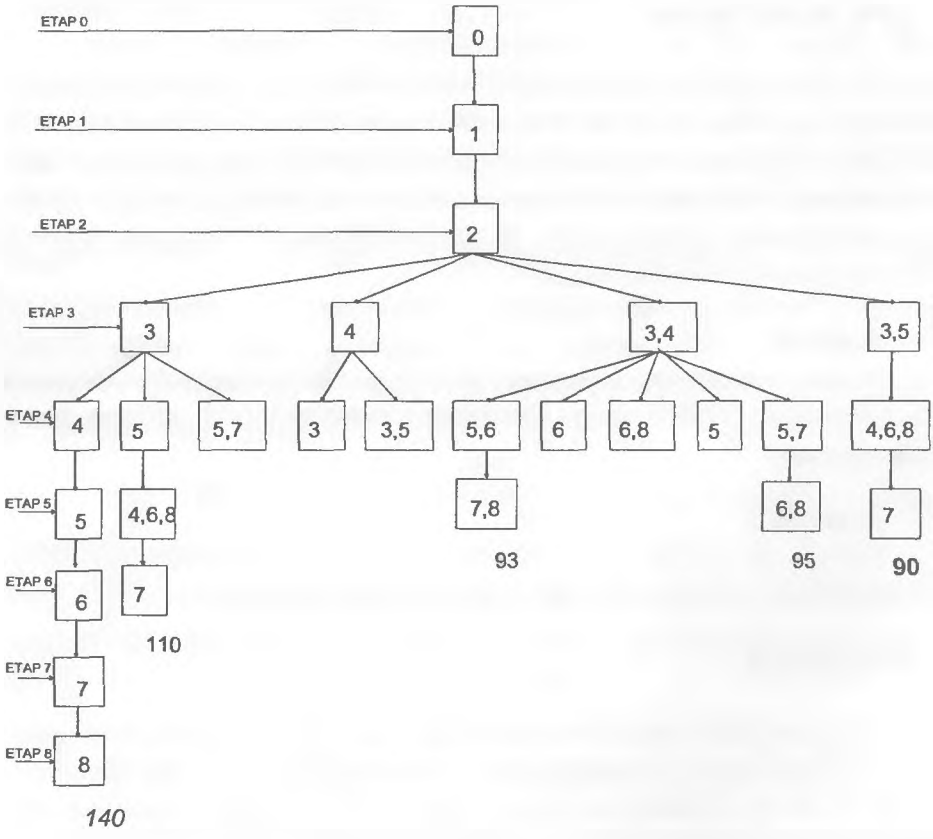
4. Przykład

Jako przykład rozpatrzono teoriogrowy opis procesu montażu na szeregowej linii montażowej z uwzględnieniem określonej kolejności operacji montażowych. Dane do przykładu pokazano na rys. 2 [3]. Na rysunku tym węzły diagramu obrazują operacje. Numery operacji zapisano w węzłach, a czasy operacji – obok węzłów. Łuki przedstawiają ograniczenia kolejności wykonania operacji. Stąd łatwo już zbudować macierz kolejności G. Założono cykl montażu $c = 20$.



Rys. 2. Diagram kolejności wykonywania operacji

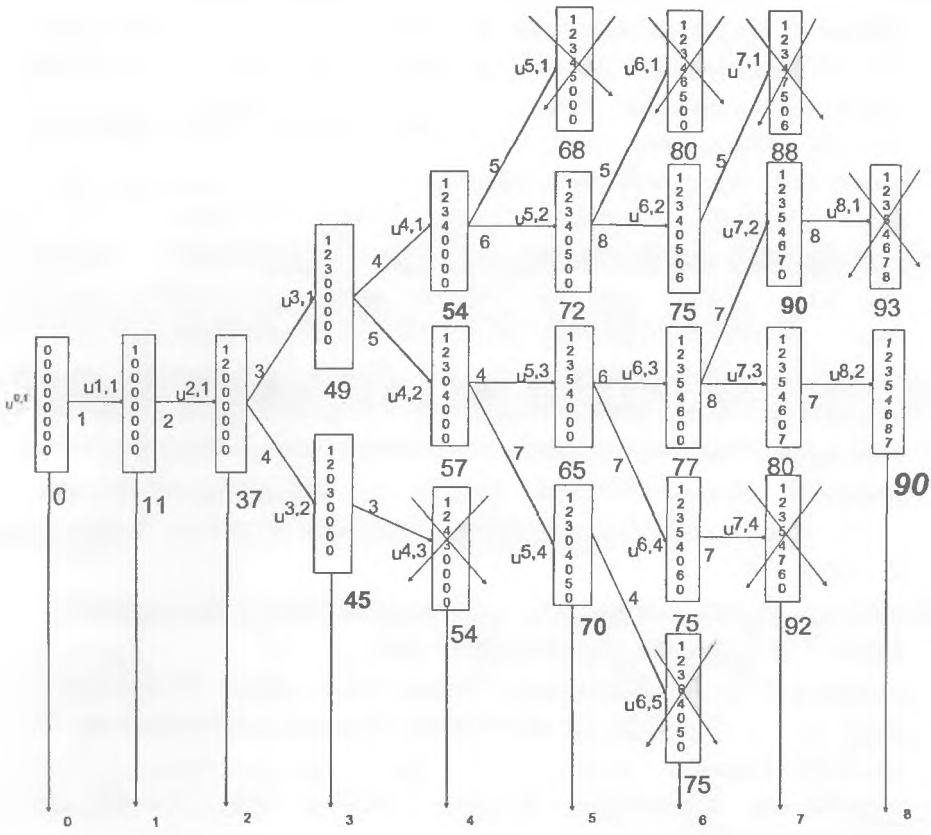
Drzewo gry z możliwymi strukturami koalicyjnymi w zależności od etapu gry przedstawiono na rys. 3.



Rys. 3. Drzewo gry

Źródło: [3].

Rozwiązując omawiany przykład za pomocą algorytmu szeregowania, otrzymano sieć stanów pokazaną na rys. 4.



Rys. 4. Sieć stanów algorytmu szeregowania

Analizując sieć widzimy, że w końcowym 8 etapie decyzyjnym znaleziono jedną optymalną sekwencję wykonywania operacji: (1,2,3,5,4,6,8,7) z minimalnym czasem 90. Biorąc pod uwagę cykl $c = 20$, natychmiast znajdujemy optymalne podzbiory operacji w postaci: $\Omega_1 = \{1\}$, $\Omega_2 = \{2\}$, $\Omega_3 = \{3,5\}$, $\Omega_4 = \{4,6,8\}$, $\Omega_5 = \{7\}$. W odpowiedniej n-osobowej grze koalicyjnej zbiory te stanowią koalicje wieloetapowej gry omówionej wcześniej.

Literatura

1. Fudenberg D., Tirole J.: Game Theory. The MIT Press, Cambridge 1991 (Fourth Printing 1995).
2. Gillies D.: Solutions to General Non-Zero-Sum Games. In: Contributions to the Theory of Games, Volume IV. Annals of Mathematics Studies, 40, Princeton University Press, Princeton 1959, pp. 47-85.

3. Kałuski J.: Game-Theoretical Model of the Assembly Line Balancing Problem. In: Proceedings of the Seventh International Symposium on Dynamic Games and Applications, Vol. 1, December 16-18, Shanon Village Center, Kanagawa, Japan 1996, pp. 417-427.
4. Krutov B.P., Novikova N.M.: Game-Theoretical Analysis of Multilevels Dynamic Hierarchic of Automatic Control System. Vychislitelnyi Centr AN SSSR, Moskva 1989 (in Russian).
5. Luce R.D., Raiffa H.: Games and Decision. John Wiley and Sons, New York 1957.
6. Marecki F.: Modele matematyczne i algorytmy alokacji operacji i zasobów na linii montażowej. Zeszyty Naukowe Politechniki Śląskiej, nr 868, Seria: AUTOMATYKA, Gliwice 1986.
7. Marecki F.: Control of Discrete Process. 5-th International Conference on Control System and Computer Science. Politechnical Institute of Bucharest, Bucharest 1983.
8. Neumann J. von, Morgenstern O.: Theory of Games and Economic Behavior. John Wiley and Sons, New York 1944.
9. Novikova N.M.: The Hierarchical *N*-Person Game with Coalitions Structure. In: Vestnik MGU, seria: „Vychislitelnaya Matematika i Kibernetika" 1979, No. 2, (in Russian).
10. Osborne M.J., Rubinstein A.: A Course in Game Theory. The MIT Press, Cambridge 1994 (Second Printing, 1995).
11. Owen G.: Game Theory. Academic Press, New York 1982 (Second Edition).
12. Shapley L.S., Shubik M.: Solutions of *N*-Person Games with Ordinal Utilities. „Economica 21" 1953, pp. 348-349.

MULTI-STAGE PROGRAMMING AND N-PERSON GAMES WITH COALITIONS

Summary

In the paper an n-stage decision process has been used to model an *n*-person game with coalitions. As an *n*-stage decision process the multi-stage programming method has been considered. The multi-stage programming method bases on the ideas of branch and bound and dynamic programming methods. The fundamental construction elements of these programming algorithms are: a state of decision process, a value of the state, state generation procedures and rules of elimination of the unperspective states.

Agnieszka Kowalska-Styczeń

MODEL UWZGLĘDNIAJĄCY PREDYSPOZYCJE PRZEDSIĘBIORSTWA DO INTERAKCJI OPARTY NA AUTOMATACH KOMÓRKOWYCH

Wstęp

W działalności przedsiębiorstw mamy do czynienia ze współpracą i interakcjami między nimi. Wiele przedsiębiorstw realizuje wspólne przedsięwzięcia w przekonaniu, że wymiana wiedzy i umiejętności przyniesie wzajemne korzyści.

W niniejszym artykule przedsiębiorstwo decyduje, czy wejść w interakcję z innym przedsiębiorstwem opierając się na wiedzy zdobytej o nim poprzez wcześniejszą interakcję oraz korzystając z informacji dodatkowych dostępnych na rynku dotyczących tego przedsiębiorstwa. Są to sygnały, które dane przedsiębiorstwo wysyła tworząc swój wizerunek na rynku.

Głównym komponentem wizerunku i reputacji przedsiębiorstwa jest zdaniem konsumentów reklama [3]. Konsumenci uważają również, że kreowanie wizerunku jest dla przedsiębiorstw konieczne, jeżeli chcą utrzymać się na konkurencyjnym rynku [3].

W przedstawionym w artykule modelu wykorzystano automaty komórkowe, o których szerzej w pracach [4; 6].

Automat komórkowy [4] to obiekt matematyczny składający się z :

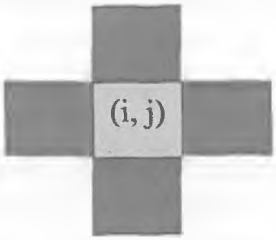
- sieci komórek {i} przestrzeni D-wymiarowej,
- zbioru stanów pojedynczej komórki $\{s_i\}$, zawierającego k elementów,
- reguły F określającej stan komórki w chwili t+1 w zależności od stanu w chwili t tej komórki i komórek ją otaczających $s_i(t+1) = F(\{s_j(t)\})$, $j \in O(i)$, gdzie O(i) jest otoczeniem i-tej komórki.

Ważne parametry dla automatu komórkowego to wymiar sieci D i ilość stanów pojedynczej komórki k.

1. Model uwzględniający predyspozycje przedsiębiorstw do interakcji

W rozważanym modelu przedsiębiorstwa wchodzące ze sobą w interakcje mogą wysyłać różne sygnały, które wskazują na ich przypuszczalne zachowanie podczas interakcji (może to być reklama tworząca wizerunek danego przedsiębiorstwa na rynku). Symulacje interakcji między przedsiębiorstwami oparto na modelu wyboru zachowania na podstawie wcześniejszych zachowań innych osób, który można znaleźć w pracy [2].

W modelu wykorzystano automaty komórkowe dwuwymiarowe z otoczeniem von Neumanna. Otoczenie takie przedstawia rys. 1.



Rys. 1. Otoczenie von Neumanna komórki (i, j) automatu dwuwymiarowego

Symulacje przeprowadzono wykorzystując funkcje i procedury dostępne w programie Mathematica 4.0 [1].

W modelu jest używana kwadratowa krata n na n, dana jest również gęstość p populacji przedsiębiorstw (zwanych dalej jednostkami) zajmujących komórki kraty. Wśród populacji występują „pozytywni" (przedsiębiorstwa, z którymi współpraca przynosi korzyści) oraz „negatywni" (przedsiębiorstwa, z którymi współpraca przynosi straty). Symulacja obejmuje t etapów.

Kratę zapełniamy w następujący sposób:

- wartość pustej komórki wynosi 0,
- wartość komórki zajętej składa się z czterech elementów $\{a, b, c, d\}$.

Elementy komórki zajętej oznaczają kolejno:

- a – kierunek, w którym jest zwrócona jednostka, wartość całkowitą od 1 do 4, wskazującą na kierunek (odpowiednio wschód, zachód, północ, południe),
- b – zachowanie przedsiębiorstwa przyjmujące wartość 0 lub 1 (0 wskazuje, że przedsiębiorstwo jest negatywne, 1 – pozytywne),
- c – natężenie sygnału emitowanego przez przedsiębiorstwo, $0 < c < 1$ (im wyższa wartość c, tym bardziej prawdopodobne jest, że dojdzie do interakcji),
- d – liczbę całkowitą wskazującą na poziom informacji (suma „wypłat" z poprzednich etapów).

W każdym wypadku jednostka otrzymuje „wypłatę" pozytywną (nagrodę) lub negatywną (karę). Macierz wypłat przedstawia tabela 1.

Do opisu macierzy wypłat wykorzystano klasyczny Dylemat Więźnia [5].

Tabela 1

Macierz wypłat dla każdego przedsiębiorstwa, które ma wybór między postępowaniem pozytywnym a negatywnym

	Pozytywne	Negatywne
Pozytywne	R R	T S
Negatywne	S T	P P

Według tej macierzy wypłat interakcja między przedsiębiorstwami może przebiegać na 4 możliwe sposoby:

- obustronna wygrana, w przypadku gdy oba przedsiębiorstwa są pozytywne i każdy dostaje nagrodę R,
- jedna wygrana i jedna przegrana, gdy w interakcję wchodzi przedsiębiorstwo pozytywne i negatywne; pozytywne otrzymuje wypłatę S („wypłatę naiwnego"), a przedsiębiorstwo negatywne wypłatę T („pokusę"),
- obie przegrane, czyli dwa przedsiębiorstwa negatywne wchodzą ze sobą w interakcję i każde otrzymuje karę P.

Jeśli nie dochodzi do interakcji, jednostka ponosi koszt uchylenia się od interakcji W.

W początkowej konfiguracji jednostki są przypadkowo rozmieszczone na kracie i zwrócone w przypadkowych kierunkach, posiadają zerowe zakresy posiadanych informacji, a sygnały mają przypadkowo wybrane wartości natężenia z przedziału (0, 1).

Podczas każdego z etapów:

- każda jednostka zwrócona w stronę innej jednostki w przyległej komórce analizuje sygnały wysyłane z tej jednostki i na podstawie ich natężenia decyduje, czy powinna wejść w interakcje, czy nie (im silniejszy sygnał, tym większe prawdopodobieństwo interakcji),
- zwrócone ku sobie jednostki, które zdecydowały się na interakcję wchodzą w interakcję, pozostałe nie,
- każda jednostka przemieszcza się do wolnej komórki z otoczenia najbliższego sąsiada, w którego stronę jest zwrócona.

W celu przeprowadzenia symulacji podaje się następujące wartości:

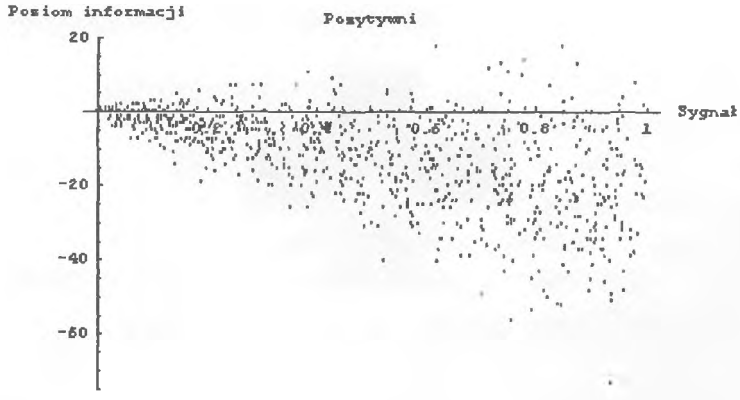
- n – rozmiar kraty,
- p – gęstość zaludnienia,

- $\{P, R, S, T, W\}$ – wartości przypisanych atrybutów,
- t – liczba etapów.

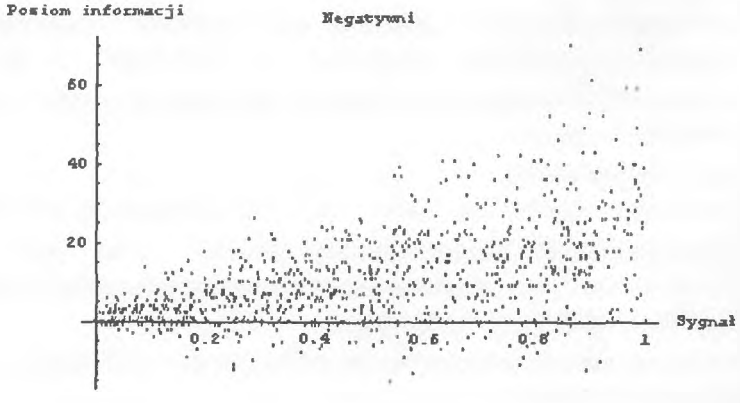
Przyjęto następujące wartości:

- $n = 50$,
- p = 70% (krata jest zaludniona przez tyle samo pozytywnych przedsiębiorstw, co negatywnych),
- $\{P, R, S, T, W\}$ wynoszą odpowiednio {-1, 1, -2, 2, 0},
- $t = 500$.

a)



b)



Rys. 2. Powyższe wykresy złożone z pojedynczych rozproszonych punktów pokazują rozkład posiadanych informacji po ostatnim etapie symulacji dla przedsiębiorstw pozytywnych (rys. a) i negatywnych (rys. b). Każdy punkt na wykresie oznacza wartość sygnału dla danej jednostki i poziom posiadanych informacji

Źródło: Wynik symulacji przeprowadzonej w programie Mathematica 4.0.

Wnioski:

- niskie wartości wysyłanego sygnału dają niewielki poziom posiadanych informacji,
- w miarę jak wartość sygnału jednostki rośnie, zwiększa się znaczenie jej zachowania.

Oznacza to, że jeżeli ilość interakcji przedsiębiorstwa z innymi przedsiębiorstwami jest niewielka, to zakres posiadanych informacji na jego temat też jest znikomy, co zmniejsza jego szanse na interakcje. Większe wartości wysyłanego przez przedsiębiorstwo sygnału zwiększają prawdopodobieństwo interakcji. Ponadto im wyższa wartość sygnału, tym znaczenie zachowania danego przedsiębiorstwa jest większe. Wynika stąd, że znane na rynku przedsiębiorstwa mają większe szanse na interakcje, ale w interakcji muszą uważać na swe zachowanie, bo będzie ono bardziej krytycznie ocenione.

W celu weryfikacji symulacji zachowań przedsiębiorstw przeprowadzono badania w formie ankiety wśród 20 osób prowadzących małe i średnie przedsiębiorstwa.

Wyniki badań:

1. 85% przedsiębiorców bierze pod uwagę reputację przedsiębiorstwa, z którym zamierza współpracować.
2. 95% przedsiębiorców nie podejmuje dalszej współpracy w przypadku niezadowolenia z pierwszego kontaktu z danym przedsiębiorstwem.
4. 75% badanych wybiera współpracę ze znanymi przedsiębiorstwami, z którymi miało już kontakt i z kontaktu tego są zadowoleni.

Przeprowadzone badania ankietowe potwierdziły wyniki symulacji. Większość przedsiębiorców nie podejmuje współpracy z firmami, z którymi współpraca wcześniej się nie powiodła. Większą szansę na interakcję mają przedsiębiorstwa o dobrej reputacji.

Literatura

1. Drwal G., Grzymkowski R., Kapusta A., Słota D.: Mathematica 4. Wydawnictwo Pracowni Komputerowej Jacka Skalmierskiego, Gliwice 2000.
2. Gaylord R.J., D'Andria L.J.: Simulating Society. Springer-Verlag, New York 1998.
3. Kieżel E.: Konsument i jego wybory rynkowe. AE, Katowice 2002.
4. Kułakowski K.: Automaty komórkowe. AGH w Krakowie, Ośrodek Edukacji Niestacjonarnej, Kraków 2000.

5. Straffin P.D.: Teoria gier. Wydawnictwo Naukowe SCHOLAR, Warszawa 2001.
6. Wolfram S.: A New Kind of Science. Wolfram Media, Inc. 2002.

MODEL BASED ON CELLULAR AUTOMATA TAKING INTO CONSIDERATION THE READINESS OF A COMPANY TO INTERACTION

Summary

In this paper a model of interaction between companies has been described. In the considered model interacting companies can transmit various signals indicating their probable behaviour during the interaction (e.g. advertisements).

In order to perform the simulation the cellular automata are used. Simulations are performed using the functions available in the program Mathematica 4.0.

It turns out that if the number of interactions of a company with other companies is small, the range of information they have got on it is also insignificant, which lowers its chance for interaction. Higher values of the signal sent by the company increase the probability of interaction. Moreover, the higher the value of the signal is the more significant the importance of a given company behaviour is. Companies known in the market have greater chance for interaction but during the interaction, they have to be careful about their behaviour as it will be assessed in a more critical way.

The results of the simulation were confirmed by the surveys conducted among people running small and medium-size companies.

Adrianna Mastalerz-Kodzis

ENTROPIA I WYKŁADNIK HURSTA A KLASYCZNE MIARY RYZYKA

Wstęp

W literaturze przedmiotu z zakresu zarówno matematyki finansowej, jak i teorii portfelowej istnieje wiele różnych miar ryzyka. W niniejszym artykule omówiono dwie niekonwencjonalne miary: entropię i wykładnik Hursta i porównano je z klasycznymi miarami ryzyka.

Zasadniczą przyczyną podjęcia tematu jest propozycja uwzględnienia niekonwencjonalnych metod pomiaru ryzyka podczas analiz. Celem artykułu jest charakterystyka entropii jako miary ryzyka oraz porównanie prezentowanej metody z klasycznymi miarami ryzyka. Pytanie, jakie nasuwa się podczas omawiania wspomnianych miar, jest następujące: czy entropia – miara ilości informacji oraz wykładnik Hursta – miara pamięci szeregu, porządkują szeregi według rosnącego ryzyka w analogiczny sposób, jak miary klasyczne? Okazuje się, że ocena ryzyka z użyciem nieklasycznych miar daje porównywalne wyniki, co wykorzystanie miar klasycznych dla rozkładu normalnego, jednakże już w przypadku rozkładu logarytmiczno-normalnego porządek ten może być odmienny.

Artykuł składa się z trzech części. Rozdział pierwszy poświęcono entropii, jej definicji, własnościom, interpretacjom oraz przykładom obliczeń zarówno dla rozkładów dyskretnych, jak i ciągłych. W rozdziale drugim definiuje się wykładnik Hursta i jego własności, zaś rozdział trzeci ma charakter empiryczny. Na podstawie wybranych walorów notowanych na Giełdzie Papierów Wartościowych w Warszawie obliczono miary ryzyka, dokonano ich porównania i z przeprowadzonych analiz wyciągnięto wnioski.

1. Entropia jako miara ilości informacji

Pojęcie **„entropia"** występuje w wielu dziedzinach nauki, m.in. w termodynamice, teorii prawdopodobieństwa, teorii informacji, teorii układów dynamicznych (w której wyróżniamy teorię procesów stochastycznych, teorię ergodyczną i dynamikę topologiczną).

Jako pierwszy pojęcia entropii użył niemiecki fizyk zajmujący się termodynamiką – Rudolf Clausius w 1854 roku. Jednakże w niniejszym artykule skupimy się na entropii pochodzącej z teorii informacji. Teoria ta zrodziła się ponad pięćdziesiąt lat temu i została zaprezentowana w książce „Matematyczna teoria komunikacji", jako przedruk artykułów pisanych przez Clauda Elwooda Shannona w „Bell System Technical Journal".

Entropia w teorii informacji wprowadzona przez C.E. Shannona to miara nieokreśloności, chaotyczności, stopnia nieuporządkowania. Jest to miara nieokreśloności doświadczenia (próby), którego wynik nie jest jednoznaczny. Entropię określa się także jako funkcję prawdopodobieństw wyników doświadczenia. Entropia zmiennej losowej charakteryzuje niepewność (losowość) wyników a priori (przed doświadczeniem). W cybernetyce entropia to miara chaotyczności układu, tym większa, im jego stany są bardziej prawdopodobne, a tym mniejsza, im jeden ze stanów jest statystycznie wyróżniony. W fizyce to wielkość charakteryzująca stan układu ciał materialnych; im większe jest prawdopodobieństwo stanu, tym większa entropia.

W literaturze pojawia się także pojęcie entropii topologicznej dla układów dynamicznych. Zostało ono wprowadzone przez R.L. Adlera, A.G. Konheima i M.H. McAndrewa w 1965 roku. Entropia jest parametrem liczbowym układu dynamicznego charakteryzującym szybkość „mieszania" punktów przez przekształcenie.

1.1. Definicja entropii

Poniżej podano definicję Shannona entropii zmiennej losowej, zarówno w przypadku zmiennej dyskretnej, jak i ciągłej [3; 4; 5].

Entropią dyskretnej zmiennej losowej X nazywamy wielkość H(X) zdefiniowaną w poniższy sposób:

$$H(X) = -\sum_{i=1}^{n} p_i \log_a p_i \qquad (1)$$

gdzie x_i jest przyjmowane z prawdopodobieństwem p_i.

Podstawą logarytmu może być dowolna liczba dodatnia. W poniższych analizach będziemy przyjmować a = 2. Im większa wartość H(X), tym większe ryzyko, czyli możliwość wystąpienia sytuacji innej aniżeli oczekiwana. Dla zmiennej dyskretnej entropia jest dodatnia. Entropia H(X) = 0 świadczy o rozkładzie stałym (jest tylko jedna możliwa wartość zmiennej losowej), zatem nie ma nieokreśloności i ryzyka.

Można także podać interpretację entropii w zależności od prawdopodobieństwa zaistnienia poszczególnych zdarzeń losowych. Entropia jest miarą nieokreśloności, zatem im mniejsze prawdopodobieństwo zajścia danego zdarzenia, tym wyższa entropia, większe ryzyko. Im większe prawdopodobieństwo zajścia zdarzenia (większa częstość występowania określonej wartości zmiennej losowej), tym mniejsza entropia, mniejsze ryzyko. W potocznym rozumieniu entropia to rosnąca funkcja prawdopodobieństwa zajścia zdarzenia.

Entropią ciągłej zmiennej losowej X nazywamy wielkość H(X) zdefiniowaną wzorem:

$$H(X) = -\int_{-\infty}^{+\infty} f(x)\log_2 f(x)dx \tag{2}$$

gdzie f(x) jest funkcją gęstości rozkładu.

Z powyższego wzoru można wnioskować, że w przypadku zmiennej o rozkładzie jednostajnym na skończonym przedziale <a, b> entropia jest zależna jedynie od długości przedziału i wyraża się wzorem:

$$H(X) = \log_2(b-a) \tag{3}$$

dla zmiennej losowej o rozkładzie normalnym entropię liczymy następująco:

$$H(X) = \frac{1}{2}\log_2\left(2\pi e\sigma^2\right) \tag{4}$$

zaś dla rozkładu logarytmiczno-normalnego zgodnie ze wzorem:

$$H(X) = \frac{1}{2\ln 2}(2R+1) + \frac{1}{2}\log_2\left(2\pi\sigma^2\right) \tag{5}$$

gdzie R jest średnią wartością zmiennej losowej, a σ^2 jest wariancją tej zmiennej.

Entropia zmiennej losowej ciągłej może przyjmować zarówno wartości dodatnie, jak i ujemne. Im większa wartość H(X), tym większe ryzyko.

1.2. Przykłady obliczania entropii dla zmiennej dyskretnej oraz ciągłej

- Zmienna losowa dyskretna X przyjmuje wartość x z prawdopodobieństwem 1, wówczas:

$$H(X) = -1 \cdot \log_2 1 = 0$$

- Zmienna losowa dyskretna X przyjmuje wartość x_1 z prawdopodobieństwem ½ oraz x_2 także z prawdopodobieństwem ½, wówczas:

$$H(X) = -\left(\tfrac{1}{2}\log_2 \tfrac{1}{2} + \tfrac{1}{2}\log_2 \tfrac{1}{2}\right) = -\left(\log_2 \tfrac{1}{2}\right) = \log_2 \left(\tfrac{1}{2}\right)^{-1} = \log_2 2 = 1$$

- Zmienna losowa ciągła X ma rozkład jednostajny na odcinku <2, 6>, wówczas jej funkcja gęstości ma postać:

$$f(x) = \begin{cases} 0 & \text{dla} \quad x \le 2 \\ 1/4 & \text{dla} \quad x \in (2,6] \\ 0 & \text{dla} \quad x > 6 \end{cases}$$

wtedy:

$$H(X) = -\int_{-\infty}^{+\infty} f(x)\log_2 f(x)dx = -\left[0 + \int_2^6 (1/4)\log_2(1/4)dx + 0\right] = -\int_2^6 (1/4)\log_2(1/4)dx =$$

$$= \left[-1/4 \cdot \log_2(1/4) \cdot x\right]_2^6 = -1/4 \cdot \log_2(1/4)(6-2) = -\log_2(1/4) = \log_2(1/4)^{-1} = \log_2 4 =$$

wzór 3

$$= \log_2(6-2) = 2$$

2. Wykładnik Hursta jako miara pamięci szeregu

Miarą zmienności szeregu jest wykładnik Hursta H [2]. Wykładnik bada długość pamięci w szeregu, czyli informuje, w jakim stopniu obserwacje z przeszłości mają wpływ na wartości teraźniejsze w szeregu. Jego wartość zawsze zawiera się w przedziale (0,1), im większa, tym dłuższa pamięć, mniejsze ryzyko zmiany. Szereg o wykładniku bliskim jedynki jest „gładki", występują małe zmiany jego wartości, a występujące zmiany mają ten sam kierunek, tzn. po wzroście wartości następuje z prawdopodobieństwem H kolejny wzrost, po spadku – spadek. Szereg o małym H jest bardzo „poszarpany", z prawdopodobieństwem (1-H) po wzroście wartości nastąpi jej spadek i na odwrót. W przypadku większości szeregów giełdowych wykładnik ten mieści się w przedziale (0,4; 0,7). Do poniższych obliczeń wykładnika Hursta użyto metody przeskalowanego zakresu opisanej np. w pracy [2].

3. Porównanie wybranych miar ryzyka na podstawie notowań cen akcji i indeksów GPW w Warszawie

Poniżej dla wybranych szeregów notowanych na Giełdzie Papierów Wartościowych w Warszawie wyznaczono klasyczne miary ryzyka oraz wykładnik Hursta i entropię. Z uzyskanych wyników wyciągnięto wnioski.

3.1. Analiza empiryczna wybranych indeksów giełdowych

Do badań posłużyły notowania wybranych indeksów GPW w Warszawie w okresie 31.12.1997-06.02.2004.

Obliczono kolejno:

- oczekiwaną stopę zwrotu indeksu R,
- wariancję stóp zwrotu σ^2,
- entropię H(X) indeksu przy założeniu, że rozkład stóp zwrotu jest rozkładem normalnym,
- H(X) indeksu, zakładając, iż rozkład stóp zwrotu jest rozkładem logarytmiczno-normalnym.

Wyniki obliczeń przedstawiono w tabeli 1, gdzie uporządkowano indeksy według rosnącej wariancji.

Tabela 1

Uporządkowanie indeksów według rosnącej wariancji

INDEKS	Oczekiwana stopa zwrotu R	Wariancja σ^2	H(X) rozkład normalny	H(X) rozkład logarytmiczno-normalny
WIRR	9,6653E-05	0,000187134	-4,144725613	-4,144586172
MIDWIG	0,00036574	0,000196882	-4,108094618	-4,107566969
WIG SPOŻYWCZY	0,00028424	0,000242407	-3,958044463	-3,957634389
WIG - PL	0,00041331	0,000269132	-3,882602472	-3,882006194
WIG	0,00041568	0,000269155	-3,882541164	-3,881941467
WIG BUDOWNICTWO	0,00012227	0,000283018	-3,846313076	-3,846136672
WIG BANKI	0,00075874	0,000302571	-3,798123169	-3,797028534
WIG 20	0,000274187	0,00040812	-3,582264204	-3,581868635
WIG TELEKOMUNIKACJA	9,9631E-06	0,000688211	-3,205335113	-3,205320739
NIF	-7,1359E-05	0,000708299	-3,184581258	-3,184684207

Z przeprowadzonych dla rozpatrywanego okresu badań empirycznych można wyciągnąć następujące wnioski:

- kolejność indeksów ze względu na rosnącą wariancję jest dokładnie taka sama, jak w przypadku entropii dla rozkładu normalnego (jest to konsekwencją wzoru 4; entropia przy założeniu rozkładu normalnego jest monotoniczną funkcją wariancji),
- ranking indeksów ze względu na rosnącą wariancję (w tym przykładzie) jest także analogiczny w przypadku entropii dla rozkładu logarytmiczno-normalnego (pomimo że w przypadku tego rozkładu entropia jest zależna zarówno od oczekiwanej stopy zwrotu, jak i od wariancji stóp zwrotu (wzór 5)).

3.2. Porównanie miar ryzyka wybranych akcji notowanych na GPW w Warszawie

Do badań posłużyły notowania wybranych akcji GPW w Warszawie w okresie 16.04.1991-08.06.2005.

Obliczono cztery miary ryzyka wymienione w przykładzie 3.1 oraz dodatkowo wykładnik Hursta. Wyniki obliczeń przedstawiono w tabeli 2.

Tabela 2

Uporządkowanie akcji według rosnącej wariancji

	KROSNO	ŚLĄSKA FABRYKA KABLI	PRÓCHNIK	TONSIL
Oczekiwana stopa zwrotu R	0,001632	0,001179	0,000655	0,000546
Wariancja σ^2	0,001382	0,001806	0,002708	0,004439
H(X) rozkład normalny	-2,70243	-2,50959	-2,21709	-1,86074
H(X) rozkład logarytmiczno-normalny	-2,70007	-2,50789	-2,21615	-1,85996
H – wykładnik Hursta	0,6242	0,6328	0,6185	0,6070

Z obliczonych w tabeli wielkości można wysunąć następujące wnioski:

1. Dla wybranych akcji GPW ranking według wzrastającej wariancji jest taki sam, jak w przypadku rosnącej entropii zarówno dla rozkładu normalnego stóp zwrotu, jak i dla rozkładu logarytmiczno-normalnego. Jest to konsekwencją pobranych do badań bardzo długich szeregów danych (ponad trzy tysiące). W przypadku długich szeregów rozkład stóp zwrotu aproksymuje się rozkładem normalnym.

2. Występuje ujemna korelacja pomiędzy wartością entropii a wykładnikiem Hursta, co jest zrozumiałe, bowiem im większa entropia, tym większa niepewność (większe ryzyko), a im mniejszy wykładnik Hursta, tym większe ryzyko. Zatem najmniej ryzykowną akcją wśród badanych jest KROSNO, zaś największym ryzykiem jest obarczony TONSIL.
3. W przypadku gdy rozkłady empirycznych stóp zwrotu są zbliżone do rozkładu normalnego, ryzyko mierzone wariancją i entropią (rozkładu normalnego) będzie wprowadzało ten sam porządek w grupie badanych walorów.

Wykładnik Hursta jest liczony w tym przypadku dla całego badanego okresu. Zakładamy tym samym, że w badanym okresie szereg był stacjonarny (jego własności nie zmieniały się pod wpływem czasu). Badania empiryczne dowodzą jednak, że szeregi giełdowe nie są stacjonarne, istnieje zatem konieczność badania ich własności lokalnie, nie zaś globalnie. Ponadto w literaturze przedmiotu wielokrotnie badano rozkłady empirycznych stóp zwrotu i okazuje się, iż w praktyce są one dalekie od rozkładu normalnego.

Nie w każdym przypadku ranking według rosnącej wariancji pokrywa się z rankingiem według także rosnącej entropii oraz malejącego wykładnika Hursta. Przypadek ten zaprezentowano w przykładzie 3.3.

3.3. Niestacjonarne szeregi akcji na GPW w Warszawie

Globalny wykładnik Hursta, liczony powyżej, daje informacje o zachowaniu się szeregu w całym badanym przedziale czasowym. Można zadać pytanie o stabilność w czasie tego wykładnika. Bowiem biorąc pod uwagę różne przedziały czasowe (np. kolejne lata), wykładnik ten zmienia swoją wartość. Analizę rozpoczęto od 1995 roku, aby szeregi danych miały porównywalną długość. W analogiczny sposób zmienia się wartość entropii. Wyniki obliczeń przedstawia tabela 3.

Tabela 3

Zmiany miar ryzyka w czasie

		KROSNO	PRÓCHNIK	ŚLĄSKA FABRYKA KABLI	TONSIL
1995	Oczekiwana stopa zwrotu R	0,003142732	-0,001979811	0,001471539	-0,000375874
	Wariancja σ^2	0,00151332	0,001498681	0,001689823	0,001739262
	Wykładnik Hursta	0,5268	0,4527	0,4915	0,4567
	H(X) rozkład normalny	-2,636938002	-2,643949717	-2,557360616	-2,536558927
	H(X) rozkład logarytmiczno-normalny	-2,632403998	-2,646805981	-2,555237634	-2,537101199

cd. tabeli 3

1996	Oczekiwana stopa zwrotu R	0,002042001	-0,000847612	-0,000410689	0,000987557
	Wariancja σ^2	0,000468137	0,000875964	0,00039033	0,000822648
	Wykładnik Hursta	0,4994	0,3736	0,4430	0,5423
	H(X) rozkład normalny	-3,483295153	-3,031324815	-3,614413717	-3,076623006
	H(X) rozkład logarytmiczno-normalny	-3,480349167	-3,032547661	-3,615006217	-3,075198263
1997	Oczekiwana stopa zwrotu R	-0,003103267	0,003052794	0,002552028	-0,000852089
	Wariancja σ^2	0,000981608	0,00199327	0,001185633	0,001004305
	Wykładnik Hursta	0,5321	0,4615	0,6323	0, 4132
	H(X) rozkład normalny	-2,949187019	-2,43822781	-2,812967622	-2,932698115
	H(X) rozkład logarytmiczno-normalny	-2,953664087	-2,43382356	-2,809285823	-2,933927419
1998	Oczekiwana stopa zwrotu R	-0,000799476	-0,001918049	-0,003750932	-0,003119846
	Wariancja σ^2	0,001337369	0,000952266	0,001870104	0,001830788
	Wykładnik Hursta	0,5670	0,2972	0,7057	0,6007
	H(X) rozkład normalny	-2,726097611	-2,971078394	-2,484237174	-2,499564204
	H(X) rozkład logarytmiczno-normalny	-2,72725101	-2,973845554	-2,489648625	-2,50406519
1999	Oczekiwana stopa zwrotu R	0,003267256	-0,00246793	0,000662172	-0,000222316
	Wariancja σ^2	0,000692958	0,001290531	0,000579887	0,001002907
	Wykładnik Hursta	0,5740	0,5807	0,6350	0,6445
	H(X) rozkład normalny	-3,20037619	-2,751814259	-3,328875298	-2,933702513
	H(X) rozkład logarytmiczno-normalny	-3,195662536	-2,755374729	-3,327919986	-2,934023248

Ryzyko mierzone za pomocą wariancji daje w przypadku analizy w czasie takie same rezultaty, jak entropia. Jednakże warto zwrócić uwagę na fakt, iż w poszczególnych latach zmieniają się wielkość ryzyka oraz ranking. Na przykład w 1998 roku najbezpieczniejszą spółką był Próchnik, zaś najbardziej ryzykowną Śląska Fabryka Kabli. W 1999 roku najmniejszym ryzykiem charakteryzuje się Śląska Fabryka Kabli, zaś obarczony największym ryzykiem jest Próchnik. Szeregi giełdowe tych spółek nie są stacjonarne.

Może także wystąpić taka sytuacja, w której ranking według entropii dla rozkładu normalnego nie będzie taki sam, jak ranking w przypadku rozkładu logarytmiczno-normalnego. Taka sytuacja może mieć miejsce, gdy do badań weźmiemy krótkie szeregi czasowe, których rozkłady nie są normalne. Poniżej podano przykładowe obliczenia dla analizowanych spółek za okres 2.11.1999--29.12.1999. Z tabeli 4 wynika, że uporządkowanie różni się w przypadku badanych entropii dla różnych rozkładów.

Tabela 4

Uporządkowanie według rosnącej entropii rozkładu normalnego

	KROSNO	ŚLĄSKA FABRYKA KABLI	TONSIL	PRÓCHNIK
H(X) rozkład normalny	-2,981813692	-2,84149601	-2,787433953	-2,761278436
H(X) rozkład logarytmiczno-normalny	-2,974756399	-2,840957673	-2,764834912	-2,78719735

Literatura

1. Kijima M.: Stochastic Processes with Applications to Finance. Chapman & Hall/CRC, London, New York 2003.
2. Mandelbrot B.B.: Fractals and Scaling in Finance. Discontinuity, Concentration, Risk. Springer-Verlag, New York 1997.
3. Michalska E.: Entropia jako miara ryzyka decyzji inwestycyjnych. W: Modelowanie Preferencji a Ryzyko '99. Red. T. Trzaskalik. AE, Katowice 1999, s. 259-270.
4. Mynarski S.: Elementy teorii systemów i informacji. AE, Kraków 1989.
5. Sobczyk K.: Stochastyczne równania różniczkowe. Wydawnictwa Naukowo-Techniczne, Warszawa 1996.

TOPOLOGICAL ENTROPY AND HURST EXPONENT COMPARED WITH CLASSIC RISK MEASUREMENT

Summary

In the literature covering the fields of financial mathematics, portfolio theory or econometrics there is a great number of different methods to measure the risk of the exchange securites. The article describes an unconventional method of measuring risk – entropy and Hurst exponent. It shows that risk measurement using entropy gives results comparable to those achieved due to the classic methods.

Anna Męczyńska

GRUPOWA OCENA EKSPERTÓW – WYBRANE HEURYSTYCZNE TECHNIKI PORZĄDKOWANIA OBIEKTÓW

Wprowadzenie

W funkcjonowaniu współczesnego przedsiębiorstwa problemem o istotnym znaczeniu jest zagadnienie podejmowania racjonalnych decyzji. Znane w praktyce techniki i narzędzia podejmowania decyzji są przystosowane do problemów programowanych, tzn. rutynowych, powtarzalnych oraz dobrze ustrukturyzowanych. Wśród współczesnych narzędzi i technik wspomagających podejmowanie decyzji dla wyżej podanych rodzajów problemów są stosowane: metody badań operacyjnych, metody symulacyjne, metody modelowe, metody informatyczne.

Jednakże w praktyce bardzo często występują decyzje podejmowane w warunkach niepewności, o charakterze nieprogramowanym, jednorazowym, zaś zadania, których one dotyczą, są słabo ustrukturyzowane. Dla tych rodzajów decyzji literatura przedmiotu zaleca stosowanie metod heurystycznych [1].

Należy zauważyć, że metody heurystyczne są jedynie uzupełnieniem metod algorytmicznych. „Sięga" się po nie wtedy, gdy zastosowanie metod algorytmicznych jest niemożliwe lub niecelowe, np. nie można uzyskać rozwiązania w zadowalającym czasie. W pewnych przypadkach istnieje wybór: albo zastosować czasochłonną, o dużym nakładzie pracy metodę algorytmiczną i wyznaczyć dokładne rozwiązanie, albo wykorzystać metodę heurystyczną (mniej pracochłonną) i ewentualnie zadowolić się przybliżonym rozwiązaniem.

Metody heurystyczne są stosowane wówczas, gdy nie dysponuje się danymi uzyskanymi na podstawie obserwacji lub nie istnieje spójna, dobrze uzasadniona oraz obiektywnie sprawdzalna teoria ustalająca korelacyjne związki

między obserwowanymi zjawiskami, albo za pomocą ekstrapolacji nie można dokonać oceny przyszłych następstw rozpatrywanych działań.

W praktyce nie zawsze można uzyskać dane z wiarygodnych źródeł. Często interesujące nas dane statystyczne mają charakter jakościowy lub niezbędne informacje są określone w sposób mało precyzyjny. Zawężenie rozważań tylko do dostępnych i precyzyjnych danych często może się okazać błędem. Wtedy można się posłużyć innymi metodami, np. metodą grupowej oceny ekspertów. Oceny ekspertów wprawdzie nie są zbyt precyzyjne, ale są wystarczająco dokładne dla większości zastosowań.

Ponieważ jakość ocen ekspertów istotnie wpływa na jakość otrzymywanych wyników, więc należy badać zespół ekspertów pod względem kompetencji [2] oraz zgodności opinii. Można przyjąć, że jeśli eksperci będą wysokiej klasy specjalistami, a ich opinie umiarkowanie zgodne, to wyniki będą wiarygodne.

W artykule opisano i porównano dwie techniki grupowej oceny ekspertów: porównywania obiektów parami (wariant z ocenami ze zbioru {0,1,2}) i względnej ważności obiektów.

1. Grupowa ocena ekspertów

Zakłada się, że na podstawie określonego kryterium Cr każdy z ekspertów potrafi określić liniowy porządek w zbiorze obiektów.

Oznaczenia:

m – liczba ekspertów biorących udział w ocenie,

n – liczba ocenianych obiektów,

c_{ij}^{k} – ocena porównania ważności i-tego oraz j-tego obiektu (z punktu widzenia kryterium Cr) wystawiona przez k-tego eksperta (i,j = 1,2,...,n; k = 1,2,...,m),

$m_{max\ j}$ – liczba ekspertów, którzy dali maksymalną ilość punktów przy ocenie j-tego obiektu,

c_{j}^{k} – ocena w punktach przyznana j-temu obiektowi przez k-tego eksperta (ekspert może przyznać od 0 do K punktów).

1.1. Wyznaczenie uogólnionej opinii ekspertów – technika porównywania obiektów parami

Uwaga: Do kroku 6 można przejść po wykonaniu kroków 1-5 dla każdego eksperta (k = 1,2,…,m).

Krok 1

Ekspert wypełnia pola następującej tabeli (tabela 1), znajdujące się nad górną przekątną. Jeśli ekspert uważa, że z punktu widzenia kryterium Cr, obiekt w i-tym wierszu jest ważniejszy od obiektu w j-tej kolumnie, to w i-tym wierszu i j-tej kolumnie wpisuje „2"; jeśli sądzi, że znaczenie obiektów jest takie samo, to w i-tym wierszu i j-tej kolumnie stawia „1"; jeśli natomiast według eksperta obiekt z i-tego wiersza ma mniejsze znaczenie niż obiekt z j-tej kolumny, to w i-tym wierszu i j-tej kolumnie wpisuje „0".

Tabela 1

Wygląd tabeli porównania obiektów parami wypełnionej przez eksperta

Nr eksperta	Obiekt 1	Obiekt 2	Obiekt 3	...	Obiekt n-1	Obiekt n
Obiekt 1		$c^k_{1,2}$	$c^k_{1,3}$	...	$c^k_{1,n-1}$	$c^k_{1,n}$
Obiekt 2			$c^k_{2,3}$	...	$c^k_{2,n-1}$	$c^k_{2,n}$
Obiekt 3				...	$c^k_{3,n-1}$	$c^k_{3,n}$
...					...	...
Obiekt n-1						$c^k_{n-1,n}$
Obiekt n						

Krok 2

Uzupełnia się tabelę porównania obiektów parami następująco: na głównej przekątnej wstawia się „1"; w i-tym wierszu i j-tej kolumnie dla i,j = 1,2,...,n oraz i > j wpisuje się wartość c^k_{ij}, gdzie:

$$c^k_{ij} = 2 - c^k_{ji} \tag{1}$$

Krok 3

Dla każdego z obiektów (i = 1,2,...,n) oblicza się ocenę C^k_i, jaką uzyskał obiekt przy ocenie k-tego eksperta (k = 1,2,...,m):

$$C^k_i = \sum_{j=1}^{n} c^k_{ij} \tag{2}$$

Krok 4

Dla każdego z obiektów (i = 1,2,...,n) wyznacza się jego rangę R^k_i, na podstawie ocen k-tego eksperta $\{C^k_i \quad i = 1,2,...,n\}$:

a. Nierosnąco porządkuje się ciąg ocen $\{C_i^k \quad i = 1,2,...,n\}$. Otrzyma się ciąg $\{O_i^k \quad i = 1,2,...,n\}$.
b. Wyznacza się rangi ocen k-tego eksperta $\tilde{R}(O_i^k)$: jeśli ciąg $\{O_i^k \quad i = 1,2,...,n\}$ jest silnie malejący, to $\tilde{R}(O_i^k) = i$; jeśli w ciągu $\{O_i^k \quad i = 1,2,...,n\}$ są wyrazy takie same, to nadaje im się rangę taką samą, równą średniej arytmetycznej rang, którą by miały, gdyby były różne.
c. Jako rangę i-tego obiektu w ocenie k-tego eksperta przyjmuje się rangę oceny, przyznanej temu obiektowi przez tego eksperta, tzn. $R_i^k = \tilde{R}(C_i^k)$.

Krok 5

Sprawdza się, czy zbiór ocen k-tego eksperta jest niesprzeczny (np. nie zachodzi warunek, że $\bigvee_{1 \le i,j,p \le n} c_{i,j}^k = 2 \wedge c_{j,p}^k = 2 \wedge c_{i,p} = 0$).

W tym celu (przy ustalonym k, k = 1,2,...,m) tworzy się pomocniczą tabelę, w której wiersze i kolumny są uporządkowane według niemałejących rang obiektów R_i^k. Tabela pomocnicza powstaje z tabeli porównywania obiektów parami przez odpowiednie przestawianie wierszy i kolumn.

Jeżeli oceny k-tego eksperta są niesprzeczne, to w tabeli pomocniczej na głównej przekątnej są same „1", nad główną przekątną są tylko „2" lub „1", natomiast pod główną przekątną występują tylko „0" lub „1"; „1" są rozmieszczone symetrycznie względem głównej przekątnej. Jeśli w i-tym wierszu i j-tej kolumnie dla $i \neq j$ znajduje się „1", to i-ty oraz j-ty wiersz powinien być identyczny, a także i-ta oraz j-ta kolumna powinny być takie same.

Jeżeli zbiór ocen jakiegoś eksperta jest sprzeczny, to należy zrezygnować z usług takiego eksperta albo poprosić go o poprawę swoich ocen tak, by oceny były niesprzeczne i powtórzyć odpowiednie kroki algorytmu.

Krok 6

Dla każdego z obiektów (i = 1,2,...,n) wyznacza się sumę rang S_i przyznanych temu obiektowi przez grupę m ekspertów według wzoru:

$$S_i = \sum_{k=1}^{m} R_i^k \tag{3}$$

Krok 7

Porządkuje się obiekty według niemalejących sum rang S_i. Uporządkowanie to wyznacza kolejność obiektów według ich znaczenia; przy czym mniejsza

suma rang odpowiada obiektowi o większym znaczeniu. Suma rang S_i odpowiadająca i-temu obiektowi reprezentuje uogólnioną opinię ekspertów o ważności i-tego obiektu.

1.2. Wyznaczenie uogólnionej opinii ekspertów – technika względnej ważności obiektów

Krok 1

Każdy ekspert ocenia obiekty z punktu widzenia kryterium Cr i każdemu z nich może przyznać od 0 do K punktów (górna granica skali K powinna być większa lub równa kilkakrotnej wielokrotności ilości obiektów, aby zapewnić możliwość przyznania różnym obiektom różnych ocen). Wyniki oceny zestawia się w macierzy $[c_{i,k}]$, gdzie $c_{i,k} = c_i^k$ dla i = 1,2,...n, k = 1,2,...m.

Krok 2

Dla każdego eksperta porządkuje się nierosnąco ciąg jego ocen (odpowiednia kolumna macierzy $[c_{i,k}]$). Otrzyma się ciąg $o_{i,k}$, gdzie k jest ustalone, a i = 1,2,...,n.

Krok 3

Dla każdego eksperta (k = 1,2,...,m) wyznacza się rangi jego ocen $\tilde{R}(o_{i,k})$: jeśli ciąg $o_{i,k}$ jest silnie malejący, to $\tilde{R}(o_{i,k}) = i$; jeśli w ciągu $o_{i,k}$ są wyrazy takie same, to nadaje im się rangę taką samą, równą średniej arytmetycznej rang, którą by miały, gdyby były różne.

Krok 4

Jako rangę i-tego obiektu w ocenie k-tego eksperta przyjmuje się rangę oceny, przyznanej temu obiektowi przez tego eksperta, tzn. $R_i^k = \tilde{R}(c_{i,k})$.

Krok 5

Dla każdego z obiektów (i = 1,2,...,n) wyznacza się sumę rang S_i przyznanych temu obiektowi przez grupę m ekspertów według wzoru 3.

Krok 6

Porządkuje się obiekty według niemalejących sum rang S_i. Uporządkowanie to wyznacza kolejność obiektów według ich znaczenia; przy czym mniejsza suma rang odpowiada obiektowi o większym znaczeniu. Suma rang S_i odpowiadająca i-temu obiektowi reprezentuje uogólnioną opinię ekspertów o ważności i-tego obiektu.

Innym wskaźnikiem uogólnionej opinii ekspertów może być, wyznaczona dla każdego i-tego obiektu (i = 1,2,...,n), średnia wartość jego oceny M_i (w punktach) dana wzorem:

$$M_i = \frac{\sum_{k=1}^{m_j} c_{ik}}{m} \tag{4}$$

Wartość M_i równa dolnej granicy skali punktowej odpowiada przypadkowi, gdy wszyscy oceniający i-ty obiekt eksperci dali najmniejszą możliwą ocenę ważności. Wartość M_i równa górnej granicy skali punktowej oznacza, że wszyscy eksperci przyznali największą możliwą ocenę. Im większa wartość M_i, tym znaczenie i-tego obiektu jest większe.

Uzupełniającym wskaźnikiem charakteryzującym uogólnioną opinię grupy ekspertów o względnej ważności obiektów jest częstość największej możliwej oceny uzyskanej przez i-ty obiekt $K_{max\,i}$ dana wzorem:

$$K_{max\,i} = \frac{m_{max\,i}}{m} \quad i = 1,...n \tag{5}$$

Wskaźnik $K_{max\,j}$ przyjmuje wartości z przedziału ⟨0; 1⟩. Charakteryzuje on znaczenie obiektu z punktu widzenia liczby przyznanych mu pierwszych miejsc.

1.3. Określenie stopnia zgodności opinii ekspertów

Przy stosowaniu metod z udziałem ekspertów pojawia się pytanie: czy uzyskane wyniki są wiarygodne? Wyniki będą wiarygodne, jeśli eksperci będą umiarkowanie zgodni w swoich opiniach. Do scharakteryzowania stopnia zgodności opinii ekspertów o względnej ważności ogółu obiektów poddanych ocenie służy współczynnik konkordancji Kendalla i Babingtona-Smitha ω (dla m > 2) [2; 3].

Jeśli wartość współczynnika konkordancji ω jest niewielka, oznacza to, że zgodność opinii jest słaba. Przyczyny mogą być różne: albo w rozpatrywanej grupie ekspertów rzeczywiście brakuje wspólnej opinii, albo wśród całego zespołu ekspertów można wyodrębnić grupy (z dużą zgodnością wewnątrzgrupową), których opinie są przeciwne.

Aby wyodrębnić grupy ekspertów, wewnątrz których jest duża zgodność opinii, można zastosować różne metody, np. metodę usuwania ocen eksperta, metodę Bartosiewicz lub metodę sumy współczynników korelacji rang Spearmana [2; 4].

2. Przykład opracowanej tabeli porównania obiektów parami

Tabela 2 przedstawia przykładową tabelę porównania obiektów parami wypełnioną przez eksperta E1 i uzupełnioną według prezentowanego wyżej algorytmu.

Te same wyniki, po uporządkowaniu obiektów według niemalejących rang, przedstawiono w tabeli 3 (tabela pomocnicza).

Tabela 2

Wyniki porównywania obiektów parami

E1	I	II	III	IV	V	Ocena	Ranga
I	1	0	0	0	0	1	5
II	2	1	0	2	1	6	2,5
III	2	2	1	2	2	9	1
IV	2	0	0	1	0	3	4
V	2	1	0	2	1	6	2,5

Tabela 3

Tabela pomocnicza (dla eksperta E1)

E1	III	II	V	IV	I	Ocena	Ranga
III	1	2	2	2	2	9	1
II	0	1	1	2	2	6	2,5
V	0	1	1	2	2	6	2,5
IV	0	0	0	1	2	3	4
I	0	0	0	0	1	1	5

W tabeli pomocniczej (dla eksperta E1) nad główną przekątną znajdują się tylko „1" i „2", a poniżej tylko „0" i „1" oraz „1" są rozmieszczone symetrycznie względem głównej przekątnej; na przecięciu wiersza i kolumny odpowiadającym obiektom II i V znajduje się „1", a wiersze (kolumny) odpowiadające tym obiektom są identyczne – świadczy to o niesprzeczności ocen eksperta E1.

3. Twierdzenie o równości rang

Oznaczenia:

n – liczba ocenianych obiektów,

$o_1, o_2, \ldots, o_n$ – oceniane obiekty,

$r_w(o_i)$ – ranga obiektu o_i określona metodą względnej ważności obiektów,

$r_p(o_i)$ – ranga obiektu o_i określona metodą porównywania obiektów parami,

$c_w(o_i)$ – ocena, jaką uzyskał obiekt o_i przy zastosowaniu metody względnej ważności obiektów,

$c_p(o_i)$ – ocena, jaką uzyskał obiekt o_i przy zastosowaniu metody porównywania obiektów parami,

$i = 1, 2, \ldots, n$.

Twierdzenie

Jeśli oceny eksperta są niesprzeczne, to rangi obiektów wyznaczone metodami względnej ważności obiektów oraz porównywania obiektów parami są równe, tzn. zachodzi:

$$\bigwedge_{i=1,2,\ldots,n} r_w(o_i) = r_p(o_i) \tag{6}$$

Dowód:

Najpierw zauważmy, że ranga obiektu jest ściśle związana z oceną, jaką ten obiekt uzyskał, oraz zachodzą następujące warunki:

$$c_w(o_i) = c_w(o_j) \Leftrightarrow r_w(o_i) = r_w(o_j) \qquad \text{dla } i,j = 1, 2, \ldots, n \; i \neq j$$

$$c_p(o_i) = c_p(o_j) \Leftrightarrow r_p(o_i) = r_p(o_j) \qquad \text{dla } i,j = 1, 2, \ldots, n \; i \neq j$$

$$c_w(o_i) < c_w(o_j) \Leftrightarrow r_w(o_i) > r_w(o_j) \qquad \text{dla } i,j = 1, 2, \ldots, n \; i \neq j$$

$$c_P(o_i) < c_P(o_j) \Leftrightarrow r_P(o_i) > r_P(o_j) \qquad \text{dla } i,j = 1, 2, \ldots, n \; i \neq j$$

Bez straty ogólności możemy założyć, że obiekty są uporządkowane nierosnąco według ich ocen uzyskanych przy zastosowaniu metody względnej ważności obiektów, tzn.:

$$\{c_w(o_1) \geq c_w(o_2) \geq \ldots \geq c_w(o_n)\} \tag{7}$$

Wtedy zachodzi:

$$\{r_w(o_1) \leq r_w(o_2) \leq \ldots \leq r_w(o_n)\} \tag{8}$$

Dowód indukcyjny twierdzenia.

Krok 1:

Niech n = 2. Mogą zajść dwa przypadki:

a) $r_w(o_1) = r_w(o_2) = 1{,}5$,

b) $r_w(o_1) = 1 \wedge r_w(o_2) = 2$.

Macierz porównań parami przedstawiono w tabelach 4 i 5, odpowiednio dla przypadku a) i b).

Tabela 4

	o_1	o_2	$c_p(o_i)$	$r_p(o_i)$
o_1	1	1	2	**1,5**
o_2	1	1	2	**1,5**

Tabela 5

	o_1	o_2	$c_p(o_i)$	$r_p(o_i)$
o_1	1	2	3	**1**
o_2	0	1	1	**2**

W obu przypadkach twierdzenie jest prawdziwe:
ad a) $r_p(o_1) = 1,5 = r_w(o_1)$, $r_p(o_2) = 1,5 = r_w(o_2)$,
ad b) $r_p(o_1) = 1 = r_w(o_1)$, $r_p(o_2) = 2 = r_w(o_2)$).

Krok 2

Założenie: twierdzenie jest prawdziwe dla n = k.
Teza: twierdzenie jest prawdziwe dla n = k + 1.

Przez $C_p^{(m)}$ oznaczmy macierz porównań parami obiektów $\{o_1,o_2,\ldots,o_m\}$. Przez $c_p^{(m)}(o_i)$ oraz $r_p^{(m)}(o_i)$ (dla $i \le m$) oznaczmy (odpowiednio) ocenę oraz rangę, jaką otrzymałby obiekt o_1, gdybyśmy zastosowali metodę porównywania parami do zbioru obiektów $\{o_1,o_2,\ldots,o_m\}$ (m „pierwszych" obiektów).

Przy przyjętych oznaczeniach zachodzi:

$$\underset{i=1,2,\ldots,m}{\wedge} c_p^{(m+1)}(o_i) = \begin{cases} c_p^{(m)}(o_i)+1 & \text{jeżeli } c_w(o_i) = c_w(o_{m+1}) \\ c_p^{(m)}(o_i)+2 & \text{jeżeli } c_w(o_i) > c_w(o_{j+1}) \end{cases} \qquad (9)$$

Ponieważ zbiór obiektów $o_1,o_2,\ldots,o_{k+1}$ jest uporządkowany i zachodzi:

$$r_w(o_1) \le r_w(o_2) \le \ldots \le r_w(o_{k+1}) \text{ oraz } c_w(o_1) \ge c_w(o_2) \ge \ldots \ge c_w(o_{k+1}),$$

więc mogą zajść dwa przypadki:

a) $c_w(o_k) > c_w(o_{k+1})$,
b) $c_w(o_k) = c_w(o_{k+1})$.

ad a)

Wtedy:

$$r_w(o_1) \le r_w(o_2) \le \ldots \le r_w(o_k) < r_w(o_{k+1}) \text{ oraz } c_w(o_1) \ge c_w(o_2) \ge \ldots \ge c_w(o_k) > c_w(o_{k+1}).$$

Macierz porównań parami wygląda następująco (tabela 6):

Tabela 6

	o_1	o_2	...	o_k	o_{k+1}
o_1	*	*	...	*	2
o_2	*	*	...	*	2
...	...	...	...	...	...
o_k	*	*	...	*	2
o_{k+1}	0	0	...	0	1

„*" oznacza, że ocena pokrywa się z odpowiednią oceną z macierzy $\mathbf{C}_p^{(k)}$.

Wtedy dla i = 1,2,...,k $c_p(o_i) = c_p^{(k+1)}(o_i) = c_p^{(k)}(o_i) + 2 \geq 2$ oraz $c_p(o_{k+1}) = 1$. Zatem:

$$\bigwedge_{i=1,2,...,k} c_p(o_{k+1}) < c_p(o_i) \tag{10}$$

stąd:

$$\bigwedge_{i=1,2,...,k} r_p(o_{k+1}) > r_p(o_k) \tag{11}$$

Z założenia indukcyjnego:

$$\bigwedge_{i=1,2,...k} r_p(o_i) = r_w(o_i) \tag{12}$$

Stąd i z (11) mamy:

$$r_p(o_{k+1}) = k+1 = r_w(o_{k+1}) \tag{13}$$

czyli:

$$\bigwedge_{i=1,2,...,k+1} r_p(o_i) = r_w(o_i)$$

ad b)

Ponieważ $c_w(o_k) = c_w(o_{k+1})$, więc niech j oznacza najmniejszy taki indeks, że $c_w(o_{k+1}) = c_w(o_j)$. Wówczas:

$$r_w(o_1) \leq r_w(o_2) \leq ... \leq r_w(o_{j-1}) < r_w(o_j) = r_w(o_{j+1}) = ... = r_w(o_{k+1})$$

oraz:

$$c_w(o_1) \geq c_w(o_2) \geq ... \geq c_w(o_{j-1}) > c_w(o_j) = c_w(o_{j+1}) = ... = c_w(o_{k+1})$$

Wtedy:

$$c_p^{(k)}(o_i) = c_p^{(k)}(o_k) \quad \text{dla} \quad k \geq i \geq j$$

Na podstawie macierzy porównań parami, przedstawionej w tabeli 7 widać, że:

$$c_p^{(k)}(o_{j-1}) > c_p^{k}(o_i) \quad \text{dla} \quad i = j, j+1, \ldots k \tag{14}$$

oraz:

$$c_p^{(k+1)}(o_i) = \begin{cases} c_p^{(k)}(o_i) + 2 & \text{dla} \quad i = 1,2,\ldots, j-1 \\ c_p^{(k)}(o_i) + 1 & \text{dla} \quad i = j, j+1, \ldots k \end{cases} \tag{15}$$

Tabela 7

	o_1	o_2	...	o_{j-1}	o_j	o_{j+1}	...	o_k	o_{k+1}
o_1	*	*	...	*	2	2	...	2	2
o_2	*	*	...	*	2	2	...	2	2
...	...	...	...	...	...	...	...	...	...
o_{j-1}	*	*	...	*	2	2	...	2	2
o_j	0	0	...	0	1	1	...	1	1
o_{j+1}	0	0	...	0	1	1	...	1	1
...	...	...	...	...	...	...	...	...	...
o_k	0	0	...	0	1	1	...	1	1
o_{k+1}	0	0	...	0	1	1	...	1	1

„*" oznacza, że ocena pokrywa się z odpowiednią oceną z macierzy $C_p^{(k)}$.

Dla $i = j, j+1, \ldots, k$ mamy:

$$c_p^{(k+1)}(o_i) = c_p^{(k)}(o_k) + 1 = c_p^{(k+1)}(o_k) = k - (j-1) + 1 = k + 2 - j \tag{16}$$

Natomiast:

$$c_p^{(k+1)}(o_{k+1}) = k + 1 - (j-1) = k + 2 - j = c_p^{k+1}(o_k) \tag{17}$$

Z założenia indukcyjnego:

$$r_p^{(k)}(o_i) = r_w^{(k)}(o_i) \quad \text{dla} \quad i = 1,2,\ldots, k$$

Stąd wobec (14) i (15) mamy, że:

$$r_p(o_i) = r_w(o_i) \quad \text{dla} \quad i = 1,2,..., j-1 \tag{18}$$

Dla i = j,j+1,...,k+1:

$$r_w(o_i) = \frac{\sum_{m=j}^{k+1} m}{k+1-(j-1)} = \frac{\frac{k+j+1}{2} * (k+j+2)}{k+j+2} = \frac{k+j+1}{2} \tag{19}$$

Wobec (14), (15), (17), dla i = j,j+1,...,k+1:

$$r_p(o_i) = r_w(o_i) \tag{20}$$

Na podstawie (18) oraz (20):

$$\bigwedge_{i=1,2,...,k+1} r_p(o_i) = r_w(o_i)$$

Wnioski

Z przedstawionych wyżej rozważań można wysnuć następujące wnioski:

1. Przy założeniach przedstawionego twierdzenia, wyniki (rangi obiektów) uzyskane z użyciem obu metod są takie same. Metodę względnej ważności obiektów można polecić jedynie w przypadkach, gdy trzeba dokonać oceny kilku obiektów. Możliwości ludzkiego umysłu są ograniczone, badacz jest w stanie „ogarnąć" i analizować równocześnie tylko kilka obiektów.
2. W przypadku gdy należy ocenić więcej obiektów, powinno się stosować metodę porównywania parami. Metoda ta jest wprawdzie bardziej czasochłonna (trzeba dokonać n*(n-1)/2 porównań), ale jest bliższa naturalnemu sposobowi dokonywania oceny przez człowieka. Dzięki wielokrotnemu porównaniu obiektu z innymi, daje możliwość kontrolowania, czy oceny eksperta są niesprzeczne.

Literatura

1. Helmer O.: Korzystanie z ocen ekspertów. Red. W. Findeisen. Analiza systemowa – podstawy i metodologia. PWN, Warszawa 1985.
2. Męczyńska A.: Metoda heurystyczna – grupowa ocena ekspertów w zastosowaniu do analizy procesów, produktów. W: Komputerowo zintegrowane zarządzanie. Red. R. Knosala. WNT, Warszawa 1999.

3. Steczkowski J., Zeliaś A.: Statystyczne metody analizy cech jakościowych. PWE, Warszawa 1981.
4. Grabiński T., Wydymus S., Zeliaś A.: Metody doboru zmiennych w modelach ekonometrycznych. PWN, Warszawa 1982.

GROUP EXPERTS' VALUATION – SELECTED HEURISTIC TECHNIQUES OF THE OBJECTS GROUPING

Summary

In the article the conditions of appealing the experts opinions have been presented. The two techniques of group experts' valuation have been described: of comparing objects in pairs and of relative importance of objects. The theorem about the equality of the objects ranks obtained via above presented techniques has been formulated and introduced.

Jerzy Mika

ROZWIĄZANIA UOGÓLNIONE WYBRANYCH ZADAŃ OPTYMALIZACJI NIELINIOWEJ

1. Metody kierunków dopuszczalnych w wypukłej optymalizacji nieliniowej

Teoria optymalizacji umożliwia efektywne poszukiwanie optymalnych rozwiązań nieliniowych zadań optymalizacyjnych wieloma konkurencyjnymi metodami. Dla rozwiązywania szczególnych zadań optymalizacji wypukłej w postaci:

$$min\{f(x)\ ;\ AX = b \wedge X \geq \Theta\}$$

podstawowe znaczenie ma tzw. metoda kierunków dopuszczalnych. Ogólna idea metody kierunków dopuszczalnych (por. np. [13]) polega na zastosowaniu iteracyjnej procedury generowania kolejnych rozwiązań dopuszczalnych rozpatrywanego zadania optymalizacyjnego, coraz lepszych w sensie kryterium optymalizacyjnego, przy czym generowanie kolejnych rozwiązań dopuszczalnych odbywa się poprzez przemieszczanie się z wcześniej ustalonych lub wyznaczonych rozwiązań dopuszczalnych:

$$X^k \in D \text{ dla } k = 0, 1, 2, \ldots$$

do następnych rozwiązań dopuszczalnych $X^{k+1} \in D$ w odpowiednio wyznaczanych kierunkach nazywanych kierunkami użytecznymi dla minimalizacji funkcji $f(X)$ lub inaczej kierunkami stosowalnymi dla minimalizacji albo też kierunkami dopuszczalnymi i stosowalnymi dla minimalizacji funkcji kryterium $f(X)$ na zbiorze rozwiązań dopuszczalnych D oraz dla rozwiązania $X^k \in D$. Tak więc kierunkiem użytecznym dla minimalizacji funkcji $f(X)$ będzie dowolny wektor $v \in R^n$, spełniający następującą formułę:

$$\bigvee_{0<\Delta\in R} \bigwedge_{\delta<\Delta} [X+\delta v \in D \wedge f(X+\delta v) < f(X)]$$

W teorii optymalizacji dowodzi się, że dla zadań optymalizacji nieliniowej rozpatrywanego typu i spełniającego przyjęte wyżej warunki dodatkowe, wektor $v \in R^n$ będzie kierunkiem użytecznym dla minimalizacji funkcji $f(X)$ na zbiorze rozwiązań dopuszczalnych D oraz dla rozwiązania $X^k \in D$, jeżeli będzie spełniona następująca koniunkcja:

$$\langle \nabla f(X^K),\ v \rangle \wedge \bigvee_{0<\Delta\in R} \bigwedge_{\delta<\Delta} X^k + \delta v \in D$$

gdzie:
$\langle \bullet, \bullet \rangle$ – symbol iloczynu skalarnego dwóch wektorów,
$\nabla f(X^k)$ – gradient funkcji w punkcie $X^k \in D$.

W związku z powyższym można wnioskować, że dla dowolnych rozwiązań dopuszczalnych $X^k \in D \wedge X \in D$ rozpatrywanego zadania optymalizacyjnego następujący wektor będący ich różnicą:

$$v_k(X) = X - X^k$$

będzie kierunkiem użytecznym dla minimalizacji wypukłej funkcji kryterium $f(X)$ na niepustym i ograniczonym zbiorze rozwiązań dopuszczalnych D oraz dla rozwiązania dopuszczalnego $X^k \in D$, wtedy i tylko wtedy, gdy spełniona będzie następująca nierówność:

$$\langle \nabla f(X^K),\ v_k(X) \rangle < 0$$

a w konsekwencji nierówność:

$$\langle \nabla f(X^K),\ X \rangle \leq \langle \nabla f(X^K),\ X^k \rangle$$

Wobec tego można zauważyć, że ostatnia z powyższych nierówności będzie spełniona w szczególności dla wektora $X \in R^n$ będącego optymalnym rozwiązaniem następującego zadania optymalizacji liniowej w postaci kanonicznej:

$$\min\{\langle \nabla f(X^K),\ X \rangle;\ X \in D\}$$

Wektor spełniający powyższy postulat można przedstawić również w postaci następującej formuły równoważnej:

$$\tilde{X}^k = \arg\ \min\{\langle \nabla f(X^K),\ X\rangle;\ X \in D\}$$

Ponieważ ponadto:

$$X^k \in D \wedge \tilde{X}^k \in D$$

i zbiór rozwiązań dopuszczalnych D jako zbiór rozwiązań odpowiedniego układu równań i nierówności liniowych jest zbiorem wypukłym, stwierdzamy, że również wszystkie punkty leżące na odcinku łączącym punkty X^k oraz $\tilde{X}^k$ należą do zbioru rozwiązań dopuszczalnych rozpatrywanego zadania optymalizacyjnego, czyli:

$$\bigwedge_{0\le\rho\le 1} X^k + \rho(\tilde{X}^k - X^k) = X^k + \rho v_k(\tilde{X}^k) \in D$$

W bezpośrednim związku z powyższym tokiem wywodu przyjmijmy również, że liczba rzeczywista:

$$0 \le \rho_k \le 1$$

będzie optymalnym dla minimalizacji wypukłej funkcji kryterium $f(X)$ na niepustym i ograniczonym zbiorze rozwiązań dopuszczalnych D, przesunięciem rozwiązania dopuszczalnego $X^k \in D$ w kierunku wektora $v_k(\tilde{X}^k)$ na odcinku łączącym punkty X^k oraz $\tilde{X}^k$, co oznacza, że:

$$\rho_k = arg \min_{0\le\rho\le 1} f(X^k + \rho(\tilde{X}^k - X^k))$$

oraz ponadto niech:

$$X^{k+1} = X^k + \rho(\tilde{X}^k - X^k)$$

to w konkluzji stwierdza się, że ostatecznie zachodzi następująca nierówność:

$$f(X^{k+1}) \le f(X^k)$$

Z wyżej przytoczonych wywodów formalnych wynika jednoznacznie, że jeżeli początkowym, startowym rozwiązaniem scharakteryzowanego tu postępowania iteracyjnego będzie:

$$X^k = X^0 \in D$$

oraz podstawiając kolejno $k = k + 1$ (dla $k = 0, 1, 2, \ldots$), można otrzymać następujący ciąg rozwiązań dopuszczalnych rozpatrywanego zadania optymalizacyjnego:

$$\{X^k\}_{k=0,1,2,\ldots}$$

dla którego będzie spełniony następujący ciąg nierówności:

$$\ldots f(X^{k+1}) \leq f(X^{\ }) \leq \ldots \leq f(X^1) \leq f(X^0)$$

Dla wygenerowanego w przedstawiony sposób ciągu rozwiązań dopuszczalnych $\{X^k\}_{k=0,1,2,\ldots}$, w teorii optymalizacji nieliniowej dowodzi się w szczególności, że dla rozpatrywanego wypukłego zadania optymalizacji nieliniowej z liniowym układem warunków ograniczających zachodzi:

$$\bigvee_{k_0 \in N} X^{k_0} = arg\ min\{f(X);\ X \in D\}$$

czyli że przy spełnieniu wszystkich przyjętych założeń dotyczących rozpatrywanego nieliniowego procesu optymalizacyjnego ciąg rozwiązań dopuszczalnych $\{X^k\}_{k=0,1,2,\ldots}$, generowany zgodnie z przedstawioną wcześniej procedurą iteracyjną, jest zbieżny w skończonej ilości kroków do poszukiwanego rozwiązania optymalnego rozpatrywanego nieliniowego zadania optymalizacyjnego lub:

$$\lim_{k\to\infty} X^k = \arg\ \min\{f(X);\ X \in D\}$$

czyli że przy spełnieniu wszystkich przyjętych założeń dotyczących rozpatrywanego nieliniowego procesu optymalizacyjnego ciąg rozwiązań dopuszczalnych $\{X^k\}_{k=0,1,2,\ldots}$, generowany zgodnie z przedstawioną wcześniej procedurą iteracyjną, jest zbieżny w nieskończonej ilości kroków do poszukiwanego rozwiązania optymalnego rozpatrywanego nieliniowego zadania optymalizacyjnego.

2. Procesy linearyzacji zadań optymalizacji wypukłej

Przedstawione w poprzedniej części niniejszego opracowania rozważania dotyczące koncepcji metody kierunków dopuszczalnych w optymalizacji nieli-

niowej pozwalają obecnie na sformułowanie konkretnej propozycji autorskiej wersji znanego algorytmu, tzw. metody Franka i Wolfe'a do wyznaczania optymalnych rozwiązań dla rozpatrywanej klasy nieliniowych zadań optymalizacyjnych w postaci kanonicznej z liniowym układem warunków ograniczających z dodatkowym założeniem, że funkcja kryterium $f(X)$ będzie na zbiorze rozwiązań dopuszczalnych D, klasy różniczkowalności C^1, czyli że funkcja ta ma ciągłe wszystkie pochodne cząstkowe rzędu pierwszego na zbiorze rozwiązań dopuszczalnych.

Krok wstępny

Wyznaczyć dowolne rozwiązanie dopuszczalne rozpatrywanego nieliniowego zadania optymalizacji wypukłej, czyli punkt $X^0 \in D$, który w dalszych procedurach będzie traktowany jako startowe rozwiązanie początkowe.

Krok 0

Wyznaczyć następujący wektor:

$$\tilde{X}^0 = \arg\ \min\{\langle \nabla f(X^0),\ X\rangle;\ X \in D\}$$

oraz w konsekwencji następny wektor w postaci:

$$X^1 = X^0 + \rho_o(\tilde{X}^0 - X^0)$$

gdzie:

$$\rho_0 = arg \min_{0 \le \rho \le 1} f\left(X^0 + \rho(\tilde{X}^0 - X^0)\right)$$

Ogólnie w proponowanej procedurze optymalizacyjnej rekomenduje się w następnej kolejności:

Krok k (dla $k \geq 1$)

Wyznaczyć następujący wektor:

$$\tilde{X}^k = \arg\ \min\{\langle \nabla f(X^k),\ X\rangle;\ X \in D\}$$

oraz w konsekwencji następny wektor w postaci:

$$X^{k+1} = X^k + \rho_k(\tilde{X}^k - X^k)$$

gdzie:

$$\rho_k = arg \min_{0 \le \rho \le 1} f\left(X^k + \rho(\widetilde{X}^k - X^k)\right)$$

Na podstawie przedstawionych wcześniej faktów z zakresu teorii metod kierunków dopuszczalnych w optymalizacji nieliniowej wnioskujemy bezpośrednio, że jeżeli dla konkretnego rozpatrywanego nieliniowego zadania optymalizacyjnego sformułowany wyżej algorytm metody Franka i Wolfe'a będzie skończony, to postępowanie algorytmiczne polegające na iteracyjnym generowaniu odpowiednich rozwiązań dopuszczalnych powinno być kontynuowane do wystąpienia dla pewnej liczby naturalnej p następującej równości:

$$X^p = X^{p-1}$$

w tym przypadku będzie dodatkowo spełniona następująca formuła:

$$\bigwedge_{q>p} X^p = X^{p-1}$$

i w konsekwencji, zgodnie z odpowiednią własnością procedur metody kierunków dopuszczalnych w optymalizacji nieliniowej, stwierdzamy, że optymalnym rozwiązaniem rozpatrywanego zadania nieliniowej optymalizacji wypukłej będzie następujący wektor:

$$X^* = X^p = \arg\ \min\{f(X);\ X \in D\}$$

W przypadku odmiennym, a więc w razie nieskończoności odpowiedniego postępowania algorytmicznego, niezbędne będzie dodatkowe ustalenie kryterium zakończenia obliczeń i wyboru przybliżonego, suboptymalnego rozwiązania rozpatrywanego zadania optymalizacyjnego. Jedną z typowych możliwości w tym zakresie jest przyjęcie postulatu zakończenia iteracyjnego procesu obliczeniowego wtedy, gdy dla pewnej liczby naturalnej N_0 i z góry zadanej liczby rzeczywistej $\varepsilon > 0$, w procedurze stosowania zmodyfikowanej wersji metody Franka i Wolfe'a zajdzie następująca nierówność:

$$\left\| X^{N_0} - X^{N_0-1} \right\| < \varepsilon$$

W tym przypadku finalnie proponuje się przyjęcie w charakterze suboptymalnego rozwiązania przybliżonego rozpatrywanego zadania optymalizacyjnego następującego wektora:

$$X^{N_0} = X^* = \arg\ \min\{f(X);\ X \in D\}$$

Powyższe rozważania prowadzą do wniosku, że w każdym kroku proponowanego postępowania algorytmicznego, aby wyznaczyć kolejny wektor będący lepszym przybliżeniem poszukiwanego rozwiązania optymalnego wyjściowego wypukłego zadania optymalizacji nieliniowej, należy rozwiązać odpowiednie pomocnicze zadanie optymalizacyjne z następującego ciągu zadań optymalizacji liniowej:

$$\{Z_k\} = \{min\{\langle \nabla f(X^k),\ X\rangle;\ X \in D\}\}, \qquad \text{dla} \qquad k = 0,\ 1,\ 2,...$$

przy czym charakterystycznym zjawiskiem towarzyszącym jest fakt, że wszystkie pomocnicze zadania optymalizacji liniowej z ciągu $\{Z_k\}$ mają ten sam zbiór rozwiązań dopuszczalnych i różnią się wyłącznie postacią analityczną liniowej funkcji kryterium.

W dalszych rozważaniach przedstawiono autorską propozycję wykorzystania i zastosowania tzw. metody uogólnionych macierzy odwrotnych w optymalizacji liniowej do efektywnej realizacji wybranych procedur wyznaczania optymalnych rozwiązań wypukłych zadań optymalizacji nieliniowej z liniowym układem warunków ograniczających.

3. Metoda wyznaczania optymalnych rozwiązań wypukłych zadań optymalizacji nieliniowej z procedurą uogólnionych macierzy odwrotnych

Rekomendowana obecnie autorska propozycja wykorzystania i zastosowania scharakteryzowanej wyżej tzw. metody uogólnionych macierzy odwrotnych w optymalizacji liniowej do efektywnej realizacji wybranych procedur wyznaczania optymalnych rozwiązań wypukłych zadań optymalizacji nieliniowej z liniowym układem warunków ograniczających będzie syntezą modyfikacji algorytmu metody Franka i Wolfe'a, przeznaczonej do wyznaczania optymalnych rozwiązań dla rozpatrywanej klasy nieliniowych zadań optymalizacyjnych i metody uogólnionych macierzy odwrotnych w optymalizacji liniowej.

Bazą teoretyczną dalszych rozważań w tym zakresie będzie konstatacja faktu, że na podstawie wcześniej przeprowadzonych rozważań zastosowanie procedury metody kierunków dopuszczalnych i modyfikacji algorytmu metody Franka i Wolfe'a (por. np. [14]) do iteracyjnego procesu poszukiwania rozwiązań optymalnych rozpatrywanej klasy wypukłych nieliniowych zadań optymalizacyjnych w postaci kanonicznej z liniowym układem warunków ograniczających, daje w kolejnych krokach następujące wyniki:

$$\hat{X}^{k-1} = arg\ min\{\langle \nabla f(X^{K-1}),\ X\rangle;\ X \in D\} \qquad \text{dla } k = 1, 2, 3, \ldots$$

lub:

$$\hat{X}^{k-1} = arg\ min\{\langle \nabla f(X^{K-1}),\ X\rangle;\ AX = b\ \wedge\ X \geq \Theta\} \quad \text{dla } k = 1, 2, 3, \ldots$$

czyli:

$$\hat{X}^{k-1} = \hat{X}^{k-1}(\hat{z}^{k}, U_0^{k-1}) \qquad \text{dla } k = 1, 2, 3, \ldots$$

i ostatecznie, zgodnie z [10]:

$$\hat{X}^{k-1} = A^{+}b + \frac{M^T}{\|M\|^2}(\hat{z}^{k-1} - CA^{+}b) + (E - A^{+}A - \frac{M^T M}{\|M\|^2})U_0^{k-1} \quad \text{dla } k = 1, 2, 3, \ldots$$

gdzie:

$$M = \nabla f(X^{k-1})(E - A^{+}A)$$

oraz:

$$\hat{z}^{k-1} = \{z^{k-1} \in R;\ \bigvee_{U^{k-1} \in R^n} X^{k-1} = X^{k-1}(z^{k-1}, U^{k-1}) \geq \Theta\}$$

Czyli $z^{k-1} \in R$ jest najmniejszą możliwą liczbą rzeczywistą spełniającą następującą formułę:

$$\bigvee_{U^{k-1} \in R^n} X^{k-1} = A^{+}b + \frac{M^T}{\|M\|^2}(z^{k-1} - CA^{+}b) + (E - A^{+}A - \frac{M^T M}{\|M\|^2})U^{k-1} \geq \Theta\}$$

gdzie:

U^{k-1} – dowolny wektor z R^n, dla którego realizuje się minimum określone powyższym wzorem.

Ponadto, zgodnie ideą i zasadami postępowania w iteracyjnej realizacji koncepcji kierunków dopuszczalnych w optymalizacji nieliniowej, kolejnym otrzymanym rozwiązaniem dopuszczalnym będzie:

$$X^k = X^{k-1} + \rho_{k-1}(\hat{X}^{k-1} - X^{k-1})$$

dla:

$$\rho_{k-1} = arg \min_{0 \le \rho \le 1} f\left(X^{k-1} + \rho(\hat{X}^k - X^{k-1})\right)$$

Na podstawie tych rezultatów oraz wcześniejszych ustaleń dotyczących otrzymanego ciągu:

$$\left\{\hat{X}^{k-1}\right\}_{k=0,1,2,\ldots}$$

kolejnych przybliżeń poszukiwanego rozwiązania optymalnego rozpatrywanego nieliniowego zadania optymalizacyjnego, a w szczególności ustaleń dotyczących zbieżności i kryteriów zakończenia postępowania iteracyjnego metody kierunków dopuszczalnych oraz algorytmu metody Franka i Wolfe'a, można obecnie stwierdzić, że jeżeli zbiór rozwiązań dopuszczalnych D rozpatrywanego zadania optymalizacji wypukłej będzie niepusty i ograniczony, funkcja kryterium *f(X)* tego zadania będzie jednomodalna, różniczkowalna oraz mająca ciągłe wszystkie pochodne cząstkowe rzędu pierwszego na zbiorze rozwiązań dopuszczalnych *D*, to:

$$\bigvee_{k_0 \in N} X^* = arg\ min\{f(X);\ X \in D\} = \hat{X}^{k_0} = \hat{X}^{k_0}(\hat{z}^{k_0}, U_0^{k_0})$$

czyli:

$$\bigvee_{k_0 \in N} X^* = arg\ min\{f(X);\ X \in D\} = A^+b + \frac{M^T}{\|M\|^2}(\hat{z}^{k_0} - CA^+b) + (E - A^+A - \frac{M^TM}{\|M\|^2})U_0^{k_0}$$

lub:

$$X^* = arg\ min\{f(X);\ X \in D\} = \lim_{k \to \infty} \hat{X}^k(\hat{z}^k, U_0^k)$$

czyli:

$$X^* = arg\ min\{f(X);\ X \in D\} = \lim_{k \to \infty}[A^+b + \frac{M^T}{\|M\|^2}(\hat{z}^{k_0} - CA^+b) + (E - A^+A - \frac{M^TM}{\|M\|^2})U]$$

Uzasadniona we wcześniejszych częściach niniejszej pracy prawdziwość powyższych faktów pozwala obecnie na przedstawienie następującej autorskiej propozycji iteracyjnego algorytmu metody Franka i Wolfe'a z procedurą metody uogólnionych macierzy odwrotnych w optymalizacji liniowej do wyznaczania optymalnych rozwiązań wypukłych zadań optymalizacji nieliniowej z liniowym układem warunków ograniczających spełniających dodatkowo wszystkie założenia przyjęte w powyższych rozważaniach.

Krok wstępny

Wyznaczyć dowolne rozwiązanie dopuszczalne $X^0 \in D$ wyjściowego nieliniowego zadania optymalizacyjnego z liniowym układem warunków ograniczających.

Krok 0

Wyznaczyć wektor:

$$\hat{X}^0 = \hat{X}^0(\hat{z}^0, U_0^0) = A^+ b + \frac{M^T}{\|M\|^2}(\hat{z}^0 - CA^+ b) + (E - A^+ A - \frac{M^T M}{\|M\|^2})U_0^0$$

dla:

$$M = \nabla f(X^0)(E - A^+ A)$$

oraz:

$$\hat{z}^0 = \{z^0 \in R; \bigvee_{U^0 \in R^n} X^0 = X^0(z^0, U^0) \geq \Theta\}$$

gdzie:

U^0 – dowolny wektor z R^n, dla którego realizuje się minimum określone powyższym wzorem.

Ponadto wyznaczyć kolejne przybliżenie poszukiwanego rozwiązania optymalnego rozpatrywanego nieliniowego zadania optymalizacyjnego:

$$X^1 = X^0 + \rho_0(\hat{X}^0 - X^0)$$

dla liczby ρ_0 spełniającej następujący warunek:

$$\rho_0 = arg \min_{0 \le \rho \le 1} f\left(X^0 + \rho(\hat{X}^0 - X^0)\right)$$

Ogólnie:

Krok k

A. Wyznaczyć wektor:

$$\hat{X}^k = \hat{X}^k(\hat{z}^k, U_0^k) = A^+ b + \frac{M^T}{\|M\|^2}(\hat{z}^k - CA^+ b) + (E - A^+ A - \frac{M^T M}{\|M\|^2})U_0^k$$

dla:

$$M = \nabla f(X^k)(E - A^+ A)$$

oraz:

$$\hat{z}^k = \{z^k \in R; \bigvee_{U^k \in R^n} X^k = X^k\left(z^k, U^k\right) \ge \Theta\}$$

gdzie:

U^k – dowolny wektor z R^n, dla którego realizuje się minimum określone powyższym wzorem.

Ponadto wyznaczyć kolejne przybliżenie poszukiwanego rozwiązania optymalnego rozpatrywanego nieliniowego zadania optymalizacyjnego:

$$X^{k+1} = X^k + \rho_k(\hat{X}^k - X^k)$$

dla liczby ρ_k spełniającej następujący warunek:

$$\rho_k = arg \min_{0 \le \rho \le 1} f\left(X^k + \rho(\hat{X}^k - X^k)\right)$$

B. Sprawdzić, czy spełniona jest równość:

$$X^k = X^{k+1}$$

a) jeżeli powyższa równość jest spełniona, to stwierdzamy, że uzyskany w punkcie A tego kroku wektor X^{k+1} jest poszukiwanym rozwiązaniem optymalnym i postępowanie iteracyjne jest zakończone, oraz:

$$X^* = X^{k+1}$$

b) jeżeli powyższa równość nie jest spełniona, to przechodzimy do następnego punktu C tego kroku.

C. Sprawdzić, czy spełniona jest nierówność:

$$\left\| X^k - X^{k+1} \right\| < \varepsilon$$

dla pewnej z góry zadanej liczby rzeczywistej $\varepsilon > 0$.

c) jeżeli powyższa nierówność jest spełniona, to stwierdzamy, że uzyskany w punkcie A tego kroku wektor X^{k+1} jest wystarczającym przybliżeniem poszukiwanego rozwiązania optymalnego i postępowanie iteracyjne jest zakończone, oraz:

$$X^* \cong X^{k+1}$$

d) jeżeli powyższa równość nie jest spełniona, to przechodzimy do następnego punktu D tego kroku.

D. Przyjąć $k = k + 1$ i przejść do realizacji czynności punktu A następnego kroku.

Powyższe algorytmiczne postępowanie iteracyjne należy kontynuować do momentu zrealizowania się podpunktu a) punktu B pewnego kroku lub podpunktu a) punktu C pewnego kroku. Tak więc powyższy algorytm jest metodą skończoną wyznaczania optymalnych rozwiązań wypukłych zadań optymalizacji nieliniowej z liniowym układem warunków ograniczających spełniających dodatkowo wszystkie założenia przyjęte wcześniej w powyższych rozważaniach.

Wnioski i rekomendacje

W przedstawionym opracowaniu zaproponowano konkretne, autorskie uogólnienia wybranych, klasycznych metod teoretycznych, które mogą być wykorzystywane do celów wspomagania podejmowania optymalnych decyzji ekonomicznych i zarządczych z wykorzystaniem niektórych idei aplikowania do tego celu metod wykorzystujących procedury uogólnionych macierzy odwrotnych w optymalizacji liniowej.

W szczególności zaprezentowano nowe, oryginalne wyniki badań mających na celu rozszerzenie klasy rozpatrywanych zadań optymalizacyjnych. Przeprowadzone analizy merytoryczne dotyczą całokształtu problematyki wy-

znaczania optymalnych rozwiązań wypukłych zadań optymalizacji nieliniowej z liniowym układem warunków ograniczających spełniających dodatkowo założenie, że funkcja kryterium rozpatrywanego zadania optymalizacyjnego będzie na zbiorze rozwiązań dopuszczalnych, klasy różniczkowalności C^1, czyli że funkcja ta ma ciągłe wszystkie pochodne cząstkowe rzędu pierwszego na zbiorze rozwiązań dopuszczalnych. Założenie to pozwoliło na efektywne wykorzystanie elementów rachunku różniczkowego w analizach merytorycznych dotyczących zadań optymalizacyjnych rozpatrywanej postaci, a w szczególności umożliwiło zastosowanie w przedmiotowym celu tzw. metody kierunków dopuszczalnych w optymalizacji nieliniowej.

Osiągnięte, konkretne wyniki analiz teoretycznych w przedmiotowym zakresie powinny w sposób istotny rozszerzać rzeczywiste możliwości praktycznej stosowalności tzw. metod uogólnionych macierzy odwrotnych do wyznaczania klasycznych rozwiązań optymalnych i tzw. uogólnionych rozwiązań optymalnych różnych typów zadań optymalizacyjnych, zwłaszcza nieliniowych z liniowym układem warunków ograniczających. W intencji autora powinny też efektywnie umożliwiać nowe i bardziej zaawansowane zastosowania rozpatrywanych teorii w procedurach wspomagania realnych procesów decyzyjnych o charakterze ekonomicznym i zarządczym.

Literatura

1. Gass S.I.: Programowanie liniowe. PWN, Warszawa 1980.
2. Martos B.: Programowanie nieliniowe. Teoria i metody. PWN, Warszawa 1983.
3. Mika J.: O pewnej wersji metody macierzy pseudoodwrotnych w programowaniu liniowym. „Przegląd Statystyczny" 1982, nr 1.
4. Mika J.: Ob odnom metodie wyczislenia obobszczennych reszeni zadacz liniejnogo programirowania i jego primienieniach w planirowani Kurzfassungen der Beitrage. Mathematik und Kybernetik in der Ökonomie VII, Halle-Wittenberg 1982.
5. Mika J.: On Determining Optimal Generalized Solution of Different Types of Linear and Nonlinear Programming Problems by Means of Pseudoinverse Matrix Method. Kurzfassungen der Beitrage, Mathematik und Kybernetik in der Ökonomie VIII, Magdeburg 1985.
6. Mika J.: O wynikach eksperymentów obliczeniowych z zastosowaniem algorytmu metody macierzy pseudoodwrotnych w programowaniu liniowym. Prace Naukowe Akademii Ekonomicznej we Wrocławiu, nr 413, Wrocław 1989.

7. Mika J.: Uogólnione rozwiązania i korekcje sprzecznych modeli optymalizacyjnych we wspomaganiu procesów podejmowania decyzji ekonomicznych. AE, Katowice 1990.
8. Mika J.: Optymalne uogólnione rozwiązania zadań programowania matematycznego z liniowym układem warunków ograniczających w procesach wspomagania podejmowania decyzji. Prace naukowe Politechniki Śląskiej w Gliwicach, Organizacja i Zarządzanie, Z. 2, Gliwice 2000.
9. Mika J.: Rozwiązania uogólnione w procesach decyzyjnych. AE, Katowice 1994.
10. Mika J.: On the Solutions of Generalized Optimization Problems with the Linear Restrictive Constrains. „Journal of Economics and Management", Vol. 1, The Karol Adamiecki University of Economics in Katowice, Katowice 2004.
11. Mika J.: Uogólnione rozwiązania optymalne w procesach wspomagania podejmowania decyzji. Dom Ekonomisty, Biuletyn Naukowo-Informacyjny, Polskie Towarzystwo Ekonomiczne w Katowicach, nr 1/32/2004, 2005.
12. Nykowski I.: Programowanie liniowe. PWE, Warszawa 1980.
13. Wit R.: Metody programowania nieliniowego. WNT, Warszawa 1986.
14. Zangwill W.I.: Programowanie nieliniowe. WNT, Warszawa 1974.

GENERALIZED SOLUTIONS OF SOME NON-LINEAR OPTIMIZATION PROBLEMS

Summary

In the paper we suggest particular, author's generalizations of the chosen classical theoretical methods to be used to the aid of the treatment of the optimal economic and managerial decisions with the use of some ideas applied to the end of methods of the generalized inverse matrices in the linear optimization.

Monika Miśkiewicz

METODA „NAJBLIŻSZYCH SĄSIADÓW" ORAZ METODA „LEM" – PORÓWNANIE EFEKTYWNOŚCI METOD PROGNOZOWANIA ZJAWISK EKONOMICZNYCH OPISANYCH ZA POMOCĄ SZEREGÓW CZASOWYCH

Wstęp

Prognozowanie zjawisk ekonomicznych jest problemem trudnym. W ciągu lat próbowano rozwiązywać go różnymi metodami, głównie statystycznymi i ekonometrycznymi. W dobie ogromnego zainteresowania teorią chaosu oraz teorią nieliniowych systemów dynamicznych, próbuje się prognozować zjawiska ekonomiczne na podstawie metod i pojęć teorii nieliniowych systemów dynamicznych. W artykule zostaną omówione dwie metody predykcji: metoda „najbliższych sąsiadów" oraz metoda LEM (Lyapunov Exponent Method). Metoda „najbliższych sąsiadów" polega na znalezieniu w zrekonstruowanej przestrzeni stanów najbliższych sąsiadów punktu poprzedzającego prognozowaną wartość w szeregu czasowym. Natomiast metoda LEM pozwala za pomocą wykładnika Lapunowa wyznaczyć prognozowaną wartość szeregu czasowego.

Celem artykułu jest wyznaczenie przyszłych wartości szeregów czasowych, utworzonych z wybranych kursów walut oraz porównanie efektywności wybranych metod predykcji pod względem dokładności wyznaczonych prognoz. W opracowaniu pod uwagę wzięto takie waluty, jak: dolar australijski AUD, korona duńska DKK, euro EUR, funt szterling GBP, jen japoński JPY, dolar amerykański USD, w okresie od 1.01.1999 roku do 30.10.2004 roku, obejmującym 1474 notowania.

1. Rekonstrukcja przestrzeni stanów

Rekonstrukcja przestrzeni stanów polega na odtworzeniu, jedynie na podstawie jednowymiarowego szeregu obserwacji, przestrzeni stanów. W 1981 roku F. Takens, wykorzystując zmienne opóźnione, zaproponował metodę rekonstrukcji zwaną metodą opóźnień. Polega ona na wykorzystaniu d-historii, które są d-elementowymi ciągami postaci:

$$x_i^d = \left(x_i, x_{i-\tau}, x_{i-2\tau}, \ldots, x_{i-(d-1)\tau}\right) \tag{1}$$

gdzie:

$$i = (d-1)\tau + 1, \ldots, n$$

d – wymiar rekonstruowanej przestrzeni (zwany również wymiarem zanurzenia),
τ – opóźnienie.

Stąd w zrekonstruowanej d-wymiarowej przestrzeni stanów element szeregu czasowego x_i jest związany z punktem postaci (1).

Takens udowodnił, że dla $d \geq 2m + 1$, gdzie m jest wymiarem atraktora, a d jest wymiarem zanurzenia, zrekonstruowana przestrzeń stanów będzie topologicznie równoważna z „oryginalną" przestrzenią. Wyboru parametru d dokonuje się zwykle metodą prób i błędów. Tradycyjnym podejściem jest obliczanie wymiaru korelacyjnego. Natomiast czas opóźnień τ można wyznaczyć za pomocą funkcji autokorelacji lub funkcji wzajemnej informacji oraz za pomocą całki korelacyjnej. Metody wyboru parametrów d i τ zostały omówione w pracy [13].

W opracowaniu czas opóźnień wyznaczano za pomocą całki korelacyjnej [13], natomiast wymiar zanurzenia metodą fałszywych sąsiadów [13].

2. Metoda „najbliższych sąsiadów"

Rozważmy szereg czasowy złożony z n obserwacji $\{x_1, x_2, \ldots, x_n\}$. Algorytm prognozowania przyszłych wartości jest następujący [2]:

1. W d-wymiarowej zrekonstruowanej przestrzeni stanów wyznaczamy J najbliższych sąsiadów punktu x_n, w sensie odległości euklidesowej (zalecana liczba najbliższych sąsiadów $J = 2(d+1)$ [1]). W zrekonstruowanej d-wymiarowej przestrzeni stanów element x_n jest związany z punktem:

$$x_n^d = \left(x_n, x_{n-\tau}, x_{n-2\tau}, \ldots, x_{n-(d-1)\tau}\right) \tag{2}$$

gdzie:
τ – opóźnienie czasowe.

2. Obliczamy sumę:

$$d_{TOT} = \sum_{i=1}^{J} d_i \tag{3}$$

gdzie $d_i = d(x_i, x_n)$ oznacza odległość między punktami x_n i x_i, $i = 1, 2, ..., J$.

3. Wyznaczamy wagę i-tego sąsiada według wzoru:

$$w_i = \frac{1}{J-1}\left(1 - \frac{d_i}{d_{TOT}}\right) \tag{4}$$

4. Wybieramy pierwsze współrzędne x_i najbliższych sąsiadów punktu x_n i na ich podstawie określamy pierwsze współrzędne ich następników x_{i+1}, $i = 1, 2, ..., J$.
5. Obliczamy prognozę dla $n + 1$ elementu jako ważoną sumę następników pierwszych współrzędnych najbliższych sąsiadów:

$$\hat{x}_{n+1} = \sum_{i=1}^{J} w_i x_{i+1} \tag{5}$$

3. Wykładniki Lapunowa [5]

Pojęcie wykładników Lapunowa zostało rozwinięte podczas charakteryzowania wrażliwości na zmianę warunków początkowych deterministycznych systemów dynamicznych.

System dynamiczny opisany równaniem:

$$x_{t+1} = f(x_t), \quad t = 0, 1, 2, ... \tag{6}$$

gdzie:
$f: X \to X, X \subset R^m$,
X – przestrzeń stanów,
$x_t, x_{i+1} \in X$,

jest wrażliwy na zmianę warunków początkowych, jeżeli istnieje $\varepsilon > 0$, takie że dla każdego $x \in X$ oraz dla każdego otoczenia U punktu x istnieje $y \in U$ oraz $n \geq 1$, takie że $\left\| f^n(x) - f^n(y) \right\| > \varepsilon$, gdzie f^n oznacza n-krotne złożenie odwzorowania f, tzn. po skończonej liczbie iteracji odległość pomiędzy dowolnymi bliskimi punktami x i y zwiększy się o więcej niż ε. Miarą tej wrażliwości systemu są wykładniki Lapunowa.

Dla systemu dynamicznego opisanego za pomocą (6) z warunkiem początkowym x_0 wykładniki Lapunowa są zdefiniowane wzorem:

$$\lambda(x_0) = \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \ln\left|f'(x_t)\right|, \text{ dla } m = 1 \tag{7}$$

oraz uogólniając na przypadek wielowymiarowy:

$$\lambda_i(x_0) = \lim_{n\to\infty} \frac{1}{n} \ln\left|\mu_i(n, x_0)\right|, \; i = 1, \ldots, m, \text{ dla } m \geq 1 \tag{8}$$

gdzie:
$\mu_i(n,x_0)$ – wartości własne macierzy $Df^n(x_0)$,
$Df^n(x_0)$ – macierz Jacobiego odwzorowania f^n.

Wielkość $\lambda(x_0)$ pokazuje (w przybliżeniu), ile razy średnio w jednej iteracji zwiększa lub zmniejsza się odległość między sąsiednimi trajektoriami. Dodatnia wartość wykładnika wskazuje na rozciąganie się przestrzeni stanów, natomiast ujemny wykładnik jest miarą zbieżności punktów przestrzeni stanów. Dla m-wymiarowego systemu dynamicznego istnieje m wykładników Lapunowa spełniających warunek $\lambda_i \geq \lambda_{i+1}$, dla $i = 1,\ldots,m-1$. Dodatnia wartość największego wykładnika Lapunowa wskazuje na wrażliwość systemu na zmianę warunków początkowych.

W latach 80. XX wieku pojawiły się różne metody pozwalające oszacować całe spektrum wykładników Lapunowa lub chociaż ten największy. W przypadku danych pochodzących z doświadczeń wyznaczenie całego spektrum nie jest możliwe. Jednak istnieje opracowany przez A. Wolfa algorytm obliczania największego wykładnika na podstawie szeregu czasowego [8]. Metoda polega na mierzeniu oddalania się sąsiadujących ze sobą punktów w zrekonstruowanej przestrzeni stanów. Najpierw wybiera się dwa sąsiednie punkty. Następnie mierzy odległość pomiędzy nimi po upływie określonego czasu. Jeśli punkty oddalają się za bardzo, znajduje się punkt zastępczy po to, aby pominąć etap kurczenia się systemu, ponieważ wykładnik Lapunowa mierzy oddalanie się punktów w przestrzeni stanów, a nie ich zbliżanie.

4. Metoda prognozowania za pomocą wykładników Lapunowa LEM [14]

Rozważmy szereg czasowy złożony z n obserwacji $\{x_1,x_2,\ldots,x_n\}$. Przy danym opóźnieniu czasowym τ i wymiarze zanurzenia d otrzymujemy punkty zrekonstruowanej przestrzeni stanów w postaci:

$$Y(1) = \left(x_1, x_{1+\tau}, x_{1+2\tau}, \ldots, x_{1+(d-1)\tau}\right)$$

$$Y(2)=\left(x_2, x_{2+\tau}, x_{2+2\tau}, \ldots, x_{2+(d-1)\tau}\right) \tag{9}$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$Y(i)=\left(x_i, x_{i+\tau}, x_{i+2\tau}, \ldots, x_{i+(d-1)\tau}\right)$$

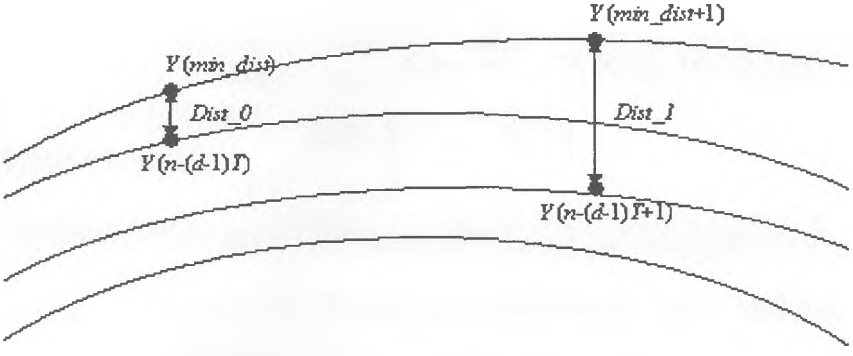
gdzie $i \in [1, n-(d-1)\tau]$. Spośród tych punktów wybieramy punkt najbliższy w sensie odległości euklidesowej punktowi $Y(n-(d-1)\tau)$ i oznaczamy przez $Y(\min\_dist)$. Minimalną odległość oznaczamy jako *Dist_0*. Jeśli *Dist_1*/*Dist_0* ulega małym zmianom w czasie ewolucji systemu, to odległość między punktami $Y(n-(d-1)\tau+1)$ i $Y(\min\_dist+1)$ wyraża się wzorem:

$$Dist\_1 = Dist\_0 \cdot 2^{K\cdot\lambda} \tag{10}$$

gdzie λ jest wykładnikiem Lapunowa, a K jest liczbą iteracji. Ponieważ

$$Y\left(n-(d-1)\tau+1\right)=\left(x_{n-(d-1)\tau+1}, x_{n-(d-1)\tau+\tau+1}, \ldots, x_{n+1}\right) \tag{11}$$

można wyznaczyć wartość x_{n+1}.



Rys. 1. Trajektorie punktów w zrekonstruowanej przestrzeni stanów

Źródło: Opracowanie własne na podstawie [14]

Algorytm prognozowania

Dany jest szereg czasowy złożony z n obserwacji $\{x_1, x_2, \ldots, x_n\}$.

1. Wybieramy opóźnienie czasowe τ oraz wymiar rekonstruowanej przestrzeni stanów d.
2. Obliczamy największy wykładnik Lapunowa λ. Jeśli $\lambda < 0$, przechodzimy do kroku 8.

3. Rekonstruujemy przestrzeń stanów dla wybranych wartości τ oraz d. Otrzymujemy $n - (d - 1)\tau$ punktów w zrekonstruowanej d-wymiarowej przestrzeni stanów.
4. Wyznaczamy punkt $Y(\min\_dist)$ położony najbliżej, w sensie odległości euklidesowej, punktu $Y(n - (d - 1)\tau)$.
5. Obliczamy odległość między punktami $Y(\min\_dist)$ oraz $Y(n - (d - 1)\tau)$.
6. Obliczamy odległość między punktami $Y(\min\_dist + 1)$ oraz $Y(n - (d - 1)\tau)$.
7. Znając współrzędne punktu $Y(\min\_dist + 1)$ w zrekonstruowanej przestrzeni stanów, prognozujemy za pomocą największego wykładnika Lapunowa kolejną wartość szeregu czasowego x_{n+1}.
8. Zmieniamy wymiar zanurzenia d i przechodzimy do kroku 2.

5. Ocena poprawności otrzymanych prognoz

Do oceny poprawności (trafności) prognozy wykorzystano bezwzględny błąd prognozy w momencie T:

$$d_T = x_T - \hat{x}_T \tag{12}$$

oraz względny błąd prognozy w momencie T:

$$D_T = \left| \frac{x_T - \hat{x}_T}{x_T} \right| \cdot 100\% \tag{13}$$

gdzie:

x_T – wartość badanej zmiennej w momencie T,

$\hat{x}_T$ – prognoza wartości zmiennej w momencie T.

Dokładność sformułowanych prognoz, w całym przedziale weryfikacji, można ocenić obliczając średni błąd prognozy ex post, który jest definiowany jako pierwiastek kwadratowy z wariancji prognozy:

$$\sigma_T = \sqrt{\frac{1}{h} \sum_{t=n+1}^{n+h} \left(x_T - \hat{x}_T\right)^2}, \quad T = n+1, ..., n+h \tag{14}$$

gdzie:

x_T – wartość badanej zmiennej w momencie T,

$\hat{x}_T$ – prognoza wartości zmiennej w momencie T,

h – liczba naturalna oznaczająca odległość okresu prognozowanego od okresu bieżącego.

Wartość średniego błędu prognozy ex post mierzy, o ile średnio odchylają się rzeczywiste realizacje zmiennej prognozowanej od obliczonych prognoz. W zastosowaniach praktycznych wygodniej posługiwać się względnym błędem średnim predykcji danym wzorem

$$V_T = \frac{\sigma_T}{\hat{x}_T} \tag{15}$$

Mała wartość średniego błędu oznacza dobre dopasowanie modelu prognostycznego do rzeczywistych realizacji zmiennej prognozowanej.

Do zbadania trafności prognoz, w całym przedziale weryfikacji, wykorzystano również współczynnik Theila wyrażony wzorem:

$$I^2 = \frac{\sum_{T=n+1}^{n+h}\left(x_T - \hat{x}_T\right)^2}{\sum_{T=n+1}^{n+h} x_T^2} = \frac{h\sigma_T^2}{\sum_{T=n+1}^{n+h} x_T^2}, \quad T = n+1, ..., n+h \tag{16}$$

W przypadku idealnie dokładnej predykcji, I^2 przybiera wartość równą zero. Pierwiastek kwadratowy z tego wyrażenia informuje, jaki był przeciętny błąd prognozy w prognozowanym okresie, bez względu na to, co było przyczyną takiego stanu rzeczy.

6. Omówienie wyników wyznaczonych prognoz

Przedstawione powyżej algorytmy prognozowania zastosowano do wyznaczenia przyszłych notowań kursów wybranych walut. Pod uwagę wzięto takie waluty, jak: dolar australijski AUD, korona duńska DKK, euro EUR, funt szterling GBP, jen japoński JPY, dolar amerykański USD, w okresie od 1.01.1999 roku do 30.10.2004 roku obejmującym 1474 notowania. Tabele 1 i 2 zawierają szczegółowe wyniki prognoz otrzymane odpowiednio metodą „najbliższych sąsiadów" oraz metodą LEM. W tabeli 3 przedstawiono średni błąd prognozy oraz współczynnik Theila wyznaczonych prognoz w całym przedziale weryfikacji dla poszczególnych metod predykcji. Pogrubioną czcionką zaznaczono wyniki, dla których błędy prognoz są najmniejsze.

Tabela 1

Wyniki prognozowania kursów walut – metoda „najbliższych sąsiadów"

	T	1	2	3	4	5
AUD	x_T	2,51	2,5223	2,5164	2,5425	2,5533
$d = 6$	$\hat{x}_T$	2,506197	0,54158	2,505841	2,526163	2,541948
$\tau = 29$	d_T	0,003803	-0,01928	0,010559	0,016337	0,011352
	D_T	0,001502	0,007574	0,004172	0,006474	0,004508
	V_T	0,00007	0,00007	0,00007	0,00007	0,000071
DKK	x_T	0,5779	0,58	0,587	0,585	0,584
$d = 6$	$\hat{x}_T$	0,57917	0,58831	0,59198	0,59027	0,58942
$\tau = 17$	d_T	-0,00127	-0,00831	-0,00498	-0,00527	-0,00542
	D_T	0,00217	0,014122	0,008524	0,009009	0,00929
	V_T	0,000052	0,000052	0,000052	0,000052	0,000052
EUR	x_T	4,2968	4,3125	4,3638	4,3492	4,3414
$d = 5$	$\hat{x}_T$	4,40886	4,4599	4,52808	4,5326	4,55788
$\tau = 9$	d_T	-0,11206	-0,1474	-0,16428	-0,1834	-0,21648
	D_T	0,025218	0,033059	0,036722	0,040833	0,047991
	V_T	0,006381	0,006359	0,006338	0,006313	0,006286
GBP	x_T	6,2119	6,2095	6,2707	6,2596	6,2484
$d = 6$	$\hat{x}_T$	6,195675	6,22244	6,223304	6,199583	6,096291
$\tau = 20$	d_T	0,016225	-0,01294	0,047396	0,060017	0,152109
	D_T	0,002613	0,002078	0,007654	0,009772	0,015356
	V_T	0,000948	0,000944	0,00095	0,000958	0,000981
JPY	x_T	3,1607	3,1711	3,1935	3,1949	3,2103
$d = 5$	$\hat{x}_T$	3,25311	3,29563	3,34041	3,3766	3,42654
$\tau = 25$	d_T	-0,09241	-0,12453	-0,14691	-0,1817	-0,21624
	D_T	2,802%	3,7625%	4,4192%	5,4373%	6,4345%
	V_T	0,007605	0,007578	0,007545	0,007505	0,007463
USD	x_T	3,396	3,3922	3,3928	3,3633	3,3397
$d = 5$	$\hat{x}_T$	3,45321	3,45599	3,4531	3,44927	3,44697
$\tau = 28$	d_T	-0,05721	-0,06379	-0,0603	-0,08597	-0,10727
	D_T	1,6566%	1,8456%	1,7462%	2,4924%	3,112%
	V_T	0,00173	0,001729	0,00173	0,001732	0,001733

Tabela 2

Wyniki prognozowania kursów walut – metoda LEM

	T	1	2	3	4	5
AUD	x_T	2,51	2,5223	2,5164	2,5425	2,5533
$d = 6$	$\hat{x}_T$	2,703063	2,703014	2,702964	2,702913	2,702859
$\tau = 29$	d_T	-0,193063	-0,180714	-0,156564	-0,160413	-0,149559
	D_T	0,071424	0,066857	0,057923	0,059348	0,055334
	V_T	0,033071	0,033072	0,033072	0,0033073	0,033074
DKK	x_T	0,5779	0,58	0,587	0,585	0,584
$d = 6$	$\hat{x}_T$	0,584526	0,585036	0,585528	0,586008	0,58648
$\tau = 17$	d_T	-0,006626	-0,005036	0,001472	-0,001008	-0,00248
	D_T	0,011336	0,008608	0,002515	0,00172	0,004229
	V_T	0,004999	0,004995	0,004991	0,004987	0,004983
EUR	x_T	4,2968	4,3125	4,3638	4,3492	4,3414
$d = 5$	$\hat{x}_T$	4,354995	4,35368	4,352001	4,349831	4,346974
$\tau = 9$	d_T	-0,05819	-0,04118	0,011799	-0,00063	-0,00557
	D_T	0,013363	0,009459	0,002711	0,000145	0,001282
	V_T	0,004756	0,004757	0,004759	0,004762	0,004765
GBP	x_T	6,2119	6,2095	6,2707	6,2596	6,2484
$d = 6$	$\hat{x}_T$	6,283835	6,293123	6,301465	6,309251	6,316689
$\tau = 20$	d_T	-0,071935	-0,083623	-0,030765	-0,049651	-0,068289
	D_T	0,011448	0,013288	0,004882	0,007869	0,010811
	V_T	0,006059	0,00605	0,006042	0,006034	0,006027
JPY	x_T	3,1607	3,1711	3,1935	3,1949	3,2103
$d = 5$	$\hat{x}_T$	3,21863	3,228465	3,239738	3,252749	3,267836
$\tau = 25$	d_T	-0,05793	-0,05736	-0,04624	-0,05785	-0,05754
	D_T	0,017998	0,017768	0,014272	0,017185	0,017607
	V_T	0,021516	0,021491	0,021416	0,021333	0,021232
USD	x_T	3,396	3,3922	3,3928	3,3633	3,3397
$d = 5$	$\hat{x}_T$	3,546991	3,465536	3,473917	3,482356	3,490996
$\tau = 28$	d_T	-0,05179	-0,09504	-0,06692	-0,07336	-0,0787
	D_T	0,014982	0,027423	0,019061	0,021065	0,022542
	V_T	0,017594	0,01755	0,017508	0,017465	0,017422

Tabela 3

Błędy prognozy w przedziale ich weryfikacji – porównanie wybranych metod predykcji

Waluta	Błędy	Metoda predykcji	
		NS	LE
AUD	σ_T	**0,000179**	0,089393
	I^2	**0,00000000497053**	0,001243565
DKK	σ_T	**0,0000305227**	0,002922
	I^2	**0,0000000027696**	0,0000251433
EUR	σ_T	0,028354	**0,020902**
	I^2	0,0000432595	**0,0000232717**
GBP	σ_T	**0,005883**	0,038073
	I^2	**0,00000089895**	0,0000372258
JPY	σ_T	**0,025081**	0,034597
	I^2	**0,0000624183**	0,0000232717
USD	σ_T	**0,005975**	0,041141
	I^2	**0,00000313066**	0,000146336

Analiza wyników predykcji wykazuje, że dla omawianych metod średnie błędy prognozy σ_T w całym przedziale weryfikacji są małe i przyjmują wartości nie większe niż 0,042. Wyjątek stanowi AUD prognozowane metodą LEM. Współczynniki rozbieżności są również zbliżone do zera, jednak metoda „najbliższych sąsiadów" daje lepsze wyniki. Wyjątek stanowi kurs EUR. Na podstawie informacji zawartych w tabelach 1-3 można stwierdzić, że dokładniejsze prognozy uzyskano metodą „najbliższych sąsiadów".

W celu poprawienia efektywności metody LEM wyznaczono prognozy dla wymiaru zanurzenia d = 2, ..., 5 oraz dla MD. MD oznacza zastosowanie metody wielowymiarowej (*multi-dimension phase space time series prediction method*), polegającej na wyznaczeniu prognoz dla poszczególnych wartości wymiaru zanurzenia d = 2, ...,5, a następnie obliczeniu średniej arytmetycznej otrzymanych prognoz. Szczegółowe wyniki, dla różnych wartości parametru d, zawierają tabele 4-10. Pogrubioną czcionką zaznaczono wyniki, dla których błędy prognoz są najmniejsze.

Tabela 4

Wyniki prognozowania kursu EUR – metoda LEM dla różnych wymiarów zanurzenia

	T	1	2	3	4	5	6	7	8	9	10
	x_T	4,2968	4,3125	4,3638	4,3492	4,3414	4,3316	4,3113	4,3093	4,3081	4,2943
$d=2$	$\hat{x}_T$	**4,315051**	**4,314526**	4,313685	4,31209	4,312392	4,314921	**4,317162**	**4,31948**	**4,322016**	**4,324869**
	d_T	**-0,01825**	**-0,00203**	0,050115	0,03711	0,029008	0,016679	**-0,00586**	**-0,01018**	**-0,01392**	**-0,03057**
	D_T	**0,00423**	**0,00047**	0,011618	0,008606	0,006727	0,003865	**0,001358**	**0,002357**	**0,00322**	**0,007068**
	V_T	**0,003839**	**0,00384**	0,003841	0,003842	0,003842	0,00384	**0,003838**	**0,003835**	**0,003833**	**0,003831**
$d=3$	$\hat{x}_T$	4,332031	4,328512	4,322094	4,32345	4,33267	4,340556	4,348417	4,356708	4,365714	4,375668
	d_T	-0,03523	-0,01601	0,041706	0,02575	0,00873	-0,00896	-0,03712	-0,04741	-0,05761	-0,08137
	D_T	0,008133	0,003699	0,00965	0,005956	0,002015	0,002063	0,008536	0,010882	0,013197	0,018596
	V_T	0,00533	0,005335	0,005343	0,005341	0,00533	0,00532	0,00531	0,0053	0,005289	0,005277
$d=4$	$\hat{x}_T$	4,352273	4,349679	4,345785	4,339532	4,327112	4,334637	4,350862	4,365649	4,380971	4,397664
	d_T	-0,05547	-0,03718	0,018015	0,009668	0,014288	-0,00304	-0,03956	-0,05635	-0,07287	-0,10337
	D_T	0,012746	0,008548	0,004145	0,002228	0,003302	0,000701	0,009093	0,012907	0,016634	0,023505
	V_T	0,006326	0,006329	0,006335	0,006344	0,006362	0,006351	0,006328	0,006306	0,006284	0,00626
$d=5$	$\hat{x}_T$	4,354995	4,35368	**4,352001**	**4,349831**	4,346974	4,343095	4,337511	40328131	4,327245	4,340705
	d_T	-0,05819	-0,04118	**0,011799**	**-0,00063**	**-0,00557**	-0,01149	-0,02621	-0,01883	-0,01914	-0,04641
	D_T	0,013363	0,009459	**0,002711**	**0,000145**	**0,001282**	0,002647	0,006043	0,004351	0,004424	0,010691
	V_T	0,004756	0,004757	**0,004759**	**0,004762**	**0,004765**	0,004769	0,004775	0,004786	0,004787	0,004772
MD	$\hat{x}_T$	4,338589	4,336599	4,333391	4,331226	4,329787	**4,333302**	4,338488	4,342492	4,348986	4,359727
	d_T	-0,04179	0,0241	0,030409	0,017974	0,011613	**-0,0017**	-0,02719	-0,03319	-0,04089	-0,06543
	D_T	0,009632	0,005557	0,007017	0,00415	0,002682	**0,000393**	0,006267	0,007644	0,009401	0,015007
	V_T	0,004484	0,004486	0,00449	0,004492	0,004493	**0,00449**	0,004484	0,00448	0,004474	0,0063

Tabela 5

Wyniki prognozowania kursu USD – metoda LEM dla różnych wymiarów zanurzenia

	T	1	2	3	4	5	6	7	8	9	10
	x_T	3,4052	3,3705	3,4077	3,409	3,4123	3,396	3,3922	3,3928	3,3633	3,3397
$d=2$	$\hat{x}_T$	**3,426776**	**3,428791**	**3,430928**	**3,433225**	**3,435718**	**3,4384444**	**3,441438**	**3,44474**	3,448392	3,452438
	d_T	**-0,02158**	**-0,05829**	**-0,02323**	**-0,02422**	**-0,02342**	**-0,02244**	**-0,04924**	**-0,05194**	-0,08505	-0,11274
	D_T	**0,006296**	**0,017001**	**0,00677**	**0,007056**	**0,006816**	**0,012344**	**0,014307**	**0,015078**	0,024676	0,032655
	V_T	**0,008941**	**0,008936**	**0,00893**	**0,008924**	**0,008918**	**0,008911**	**0,008903**	**0,008895**	0,008885	0,008875
$d=3$	$\hat{x}_T$	3,436103	3,439241	3,442556	3,446092	3,449891	3,453996	3,448449	3,463293	3,468578	3,474354
	d_T	-0,0309	-0,06874	-0,03486	-0,03709	-0,03759	-0,058	-0,06625	-0,07049	-0,10528	-0,13465
	D_T	0,008994	0,019987	0,010125	0,010763	0,010896	0,016791	0,019156	0,020354	0,030352	0,038757
	V_T	0,011025	0,011014	0,011004	0,010993	0,01098	0,010967	0,010953	0,010938	0,010921	0,010903
$d=4$	$\hat{x}_T$	3,461555	3,460178	3,458494	3,456415	3,453812	3,450485	3,446085	3,439861	**3,429119**	**3,433336**
	d_T	-0,05636	-0,08968	-0,05079	-0,04741	-0,04151	-0,05449	-0,05388	-0,4706	**-0,06582**	**-0,09364**
	D_T	0,01628	0,025917	0,014687	0,013718	0,012019	0,015791	0,015637	0,013681	**0,019194**	**0,027273**
	V_T	0,010328	0,010332	0,010337	0,010344	0,010351	0,010361	0,010375	0,010393	**0,010426**	**0,010413**
$d=5$	$\hat{x}_T$	3,546991	3,465536	3,473917	3,482356	3,490996	3,499944	3,509289	3,51911	3,529482	3,540472
	d_T	-0,05179	-0,09504	-0,06692	-0,07336	-0,0787	-0,10394	-0,11709	-0,12631	-0,16618	-0,20078
	D_T	0,014982	0,027423	0,019061	0,021065	0,22542	0,029699	0,033365	0,035893	0,047084	0,056709
	V_T	0,017594	0,01755	0,017508	0,017465	0,017422	0,017378	0,017331	0,017283	0,017232	0,017179
MD	$\hat{x}_T$	3,445356	3,448437	3,451474	3,454522	3,447604	3,460717	3,463815	3,466751	3,468893	3,475151
	d_T	-0,04016	-0,07794	-0,04377	-0,04502	-0,0453	-0,06472	-0,07162	-0,07395	-0,10559	-0, 13545
	D_T	0,011655	0,022601	0,012683	0,013177	0,013103	0,018701	0,020675	0,021332	0,03044	0,038977
	V_T	0,01166	0,01165	0,01164	0,011629	0,011619	0,011608	0,011598	0,011588	0,011581	0,01156

Tabela 6

Wyniki prognozowania kursu JPY – metoda LEM dla różnych wymiarów zanurzenia

	T	1	2	3	4	5	6	7	8	9	10
	x_T	3,1607	3,1711	3,1935	3,1949	3,2103	3,2056	3,1853	3,1715	3,1601	3,1444
$d=2$	$\hat{x}_T$	3,206215	3,209534	3,213305	3,217686	3,22845	3,228967	3,236272	3,245018	3,255512	3,268124
	d_T	-0,04551	-0,03843	-0,0198	-0,02279	-0,01254	-0,02337	-0,05097	-0,06752	-0,09541	-0,12372
	D_T	0,014196	0,011975	0,006163	0,0007082	0,003892	0,007237	0,01575	0,020807	0,029308	0,037858
	V_T	0,010236	0,010226	0,010214	0,01102	0,010183	0,010164	0,010141	0,010114	0,010081	0,010042
$d=3$	$\hat{x}_T$	**3,163674**	**3,159916**	3,15232	3,156874	3,167332	3,177032	**3,187265**	**3,198616**	**3,1211536**	**3,122646**
	d_T	**-0,00297**	**0,011184**	0,04118	0,038026	0,042968	0,028568	**-0,00197**	**-0,02112**	**-0,05144**	**-0,08206**
	D_T	**0,00094**	**0,003539**	0,013063	0,012045	0,013566	0,008992	**0,000617**	**0,006602**	**0,016016**	**0,025435**
	V_T	**0,007186**	**0,007194**	0,007212	0,007201	0,007177	0,007155	**0,0071333**	**0,007107**	**0,007079**	**0,007046**
$d=4$	$\hat{x}_T$	3,186994	3,13615	**3,200312**	**3,207492**	**3,215423**	**3,224336**	3,234463	3,246053	3,259385	3,274774
	d_T	-0,02629	-0,02252	**-0,00681**	**-0,01259**	**-0,00512**	**-0,01874**	-0,04916	-0,06855	-0,09929	-0,13037
	D_T	0,00825	0,00705	**0,002129**	**0,003926**	**0,001593**	**0,005811**	0,0152	0,021119	0,030461	0,039811
	V_T	0,010532	0,01051	**0,010488**	**0,010465**	**0,010439**	**0,01041**	0,010378	0,010341	0,010298	0,01025
$d=5$	$\hat{x}_T$	3,21863	3,228465	3,239738	3,252749	3,267836	3,285388	3,2305855	3,2329759	3,2357711	3,2390422
	d_T	-0,05793	-0,05736	-0,04624	-0,05785	-0,05754	-0,07979	-0,12056	-0,15226	-0,19761	-0,24602
	D_T	0,017998	0,017768	0,014272	0,017785	0,017607	0,024286	0,036467	0,045727	0,058853	0,072564
	V_T	0,021516	0,021491	0,021416	0,021333	0,021232	0,021118	0,020987	0,020838	0,020663	0,020464
MD	$\hat{x}_T$	3,193878	3,197883	3,201419	3,2087	3,218359	3,228931	3,40964	3,54862	3,71036	3,89946
	d_T	-0,03318	-0,02678	-0,00792	-0,0138	-0,00806	-0,02333	-0,05566	-0,07736	-0,11004	-0,14555
	D_T	0,010388	0,008375	0,002473	0,004304	0,002504	0,007226	0,017175	0,023768	0,033915	0,04424
	V_T	0,011695	0,011681	0,011668	0,011641	0,011606	0,011568	0,011526	0,011476	0,01142	0,011354

Tabela 7

Wyniki prognozowania kursu GBP – metoda LEM dla różnych wymiarów zanurzenia

	T	1	2	3	4	5	6	7	8	9	10
	x_T	6,2119	6,2095	6,2707	6,2596	6,2484	6,2245	6,2176	6,2399	6,2018	6,151
$d=2$	$\hat{x}_T$	**6,252665**	**6,251104**	6,248375	**6,241471**	**6,249432**	**6,2563,9**	**6,263208**	**6,270819**	**6,279559**	**6,289812**
	d_T	**-0,04076**	**-0,0416**	-0,022325	**-0,018129**	**-0,00103**	**-0,03181**	**-0,04561**	**-0,03092**	**-0,07776**	**-0,13881**
	D_T	**0,00652**	**0,006655**	0,003573	**0,002905**	**0,000165**	**0,005084**	**0,007282**	**0,004931**	**0,012383**	**0,022069**
	V_T	**0,0005112**	**0,005114**	0,005116	**0,005121**	**0,005115**	**0,005109**	**0,005104**	**0,005057**	**0,00509**	**0,00582**
$d=3$	$\hat{x}_T$	6,281269	6,27291	**6,269024**	6,285923	6,300291	6,315208	6,31711	6,50538	6,72392	6,98029
	d_T	-0,06937	-0,06341	**-0,01676**	-0,02632	-0,05189	-0,09071	-0,11411	-0,11064	-0,17059	-0,24703
	D_T	0,011044	0,010109	**0,000267**	0,004188	0,008236	0,014363	0,018022	0,017422	0,026771	0,03861
	V_T	0,005845	0,005858	**0,005864**	0,005838	0,005815	0,005792	0,005767	0,005738	0,005704	0,005665
$d=4$	$\hat{x}_T$	6,362323	6,359653	6,355865	6,350409	6,342352	6,329871	6,308136	6,286159	6,333938	6,371122
	d_T	-0,15042	-0,15015	-0,08516	-0,09081	-0,09395	-0,10537	-0,09054	-0,04626	-0,13214	-0,22012
	D_T	0,023643	0,02361	0,013399	0,0143	0,014813	0,016647	0,014352	0,020862	0,007359	0,03455
	V_T	0,01049	0,010495	0,010501	0,01051	0,010523	0,010544	0,010581	0,010617	0,010537	0,010476
$d=5$	$\hat{x}_T$	6,307095	6,323323	6,339223	6,355343	6,372025	6,389525	6,408059	6,427824	6,449018	6,471823
	d_T	-0,0952	-0,11382	-0,06852	-0,09574	-0,12362	-0,16502	-0,19046	-0,18792	-0,24721	-0,32082
	D_T	0,015093	0,018001	0,010809	0,015065	0,019401	0,025827	0,029722	0,029236	0,038334	0,049572
	V_T	0,01486	0,014821	0,014784	0,014747	0,014708	0,014668	0,014625	0,01458	0,014533	0,014481
MD	$\hat{x}_T$	6,300838	6,301747	6,303122	6,308287	6,316025	6,322728	6,327229	6,333835	6,358726	6,382696
	d_T	00,08894	-0,09225	-0,03242	-0,04869	-0,06763	-0,09823	-0,11018	-0,09394	-0,15683	-0,2317
	D_T	0,014115	0,014638	0,005144	0,007718	0,010707	0,015536	0,017412	0,014831	0,024679	0,036301
	V_T	0,009574	0,009573	0,00957	0,009563	0,009551	0,009541	0,009533	0,009524	0,009487	0,009451

Tabela 8

Wyniki prognozowania kursu AUD – metoda LEM dla różnych wymiarów zanurzenia

	T	1	2	3	4	5	6	7	8	9	10
	x_T	2,51	2,5223	2,5164	2,5425	2,5533	2,5355	2,5263	2,5418	2,5398	2,5296
$d=2$	$\hat{x}_T$	2,567442	2,56632	2,564728	2,562432	2,559036	2,553759	**2,544476**	**2,537346**	**2,556515**	**2,571839**
	d_T	-0,05744	-0,04402	-0,01833	-0,01993	-0,00574	-0,01826	**-0,01818**	**0,004455**	**-0,01672**	**-0,04224**
	D_T	0,022373	0,017153	0,007146	0,007779	0,002242	0,00715	**0,07143**	**0,001756**	**0,006538**	**0,016424**
	V_T	0,006517	0,00652	0,006524	0,00653	0,006538	0,006552	**0,006576**	**0,006594**	**0,006545**	**0,006506**
$d=3$	$\hat{x}_T$	**2,522226**	2,530062	2,536288	**2,54238**	**2,548753**	2,555042	2,563235	2,571705	2,581231	2,592008
	d_T	**-0,01223**	-0,00776	0,010112	**0,00012**	**0,004547**	-0,02014	-0,03693	-0,0299	-0,04143	-0,06241
	D_T	**0,004847**	0,003068	0,003987	**0,0000471**	**0,001784**	0,007881	0,014409	0,011628	0,016051	0,024077
	V_T	0,00816	0,008134	0,008114	**0,008095**	**0,008075**	0,008053	0,008029	0,008003	0,007973	0,00794
$d=4$	$\hat{x}_T$	2,531376	**2,525541**	2,524878	2,533489	2,540355	**2,546884**	2,553492	2,560396	2,567749	2,575681
	d_T	-0,02138	**-0,00324**	0,021522	0,009011	0,012945	**-0,01138**	-0,02719	-0,0186	-0,02795	-0,04608
	D_T	0,008444	**0,001283**	0,008524	0,003557	0,005096	**0,00447**	0,010649	0,007263	0,010885	0,017891
	V_T	0,006407	**0,006421**	0,006423	0,006401	0,006384	**0,006368**	0,006351	0,006334	0,006316	0,006296
$d=5$	$\hat{x}_T$	2,533675	2,526395	2,528017	2,538195	2,546828	2,555354	2,564266	2,57386	2,584372	2,596023
	d_T	-0,02367	-0,0041	0,018383	0,004305	0,006472	-0,01985	-0,03797	-0,03206	-0,04457	-0,06642
	D_T	0,009344	0,001621	0,007272	0,001696	0,0025841	0,00777	0,014806	0,012456	0,017247	0,025586
	V_T	0,008498	0,008522	0,008517	0,008483	0,008454	0,008426	0,008396	0,008365	0,008331	0,008294
MD	$\hat{x}_T$	2,53868	2,537079	**2,538478**	2,544124	2,548743	2,55291	2,556367	2,500826	2,572467	2,583888
	d_T	-0,02868	-0,01478	**0,007922**	-0,00162	0,004557	-0,01741	-0,03007	-0,01903	-0,03267	-0,05429
	D_T	0,011297	0,005825	**0,003121**	0,000638	0,001788	0,00682	0,011762	0,00743	0,012699	0,02101
	V_T	0,00663	0,006634	**0,006631**	0,006616	0,006604	0,006593	0,006584	0,006573	0,006543	0,006514

Tabela 9

Wyniki prognozowania kursu DKK – metoda LEM dla różnych wymiarów zanurzenia

	T	1	2	3	4	5	6	7	8	9	10
	x_T	0,5779	0,58	0,587	0,585	0,584	0,5827	0,5801	0,5798	0,5796	0,5777
$d=2$	$\hat{x}_T$	0,577242	0,577209	0,577173	0,577134	0,577091	0,577043	0,576991	0,576933	0,57687	0,576799
	d_T	**0,000658**	0,002791	0,009827	0,007866	0,006909	0,005657	0,003109	0,002867	0,00273	0,000901
	D_T	**0,00114**	0,004835	0,017025	0,013629	0,011972	0,009803	0,005388	0,004969	0,004733	0,001562
	V_T	0,005364	0,005364	0,005365	0,005365	0,005365	0,005366	0,005366	0,005367	0,005367	0,005368
$d=3$	$\hat{x}_T$	0,580107	0,580002	0,579885	0,579753	0,579604	0,579433	0,579234	0,578997	0,578702	0,5783
	d_T	-0,002207	**-0,00000181**	0,007115	0,005247	0,004396	0,003267	**0,000866**	**0,000803**	**0,000898**	**-0,0006**
	D_T	0,003804	**0,00000312**	0,01227	0,00905	0,007585	0,005638	**0,001494**	**0,001387**	**0,001553**	**0,001038**
	V_T	0,004164	0,004165	0,004166	0,004167	0,004168	0,004169	0,004171	0,004172	0,004174	0,004177
$d=4$	$\hat{x}_T$	0,588841	0,588683	0,588505	0,588304	0,588075	0,587815	0,587517	0,587171	0,586767	0,586287
	d_T	-0,010941	-0,008683	**-0,001505**	-0,003304	-0,004075	-0,005115	-0,007417	-0,007371	-0,007167	-0,008587
	D_T	0,018581	0,01475	**0,002557**	0,005616	0,00693	0,008702	0,012624	0,012554	0,012214	0,014646
	V_T	0,006754	0,006756	0,006758	0,00676	0,006763	0,006766	0,006769	0,006773	0,006778	0,006784
$d=5$	$\hat{x}_T$	0,583554	0,584271	0,584936	0,58557	0,586188	0,586796	0,587402	0,588009	0,588621	0,58924
	d_T	-0,005654	-0,004271	0,002064	**-0,00057**	-0,002188	-0,004096	-0,007302	-0,008209	-0,009021	-0,01154
	D_T	0,009689	0,00731	0,003529	**0,000974**	0,003732	0,006981	0,012431	0,01396	0,015325	0,019585
	V_T	0,006208	0,0062	0,006193	0,006186	0,00618	0,006174	0,006167	0,006161	0,006154	0,006148
MD	$\hat{x}_T$	0,582436	0,582541	0,582625	0,58269	0,58274	0,582772	0,582786	0,582778	0,58274	0,0582656
	d_T	-0,004536	-0,002541	0,004375	0,00231	**0,00126**	**-0,000072**	-0,002686	-0,002978	-0,00314	-0,004956
	D_T	0,007788	0,004362	0,00751	0,003964	**0,002163**	**0,000124**	0,004609	0,005109	0,005388	0,008507
	V_T	**0,003873**	**0,003872**	**0,003872**	**0,003871**	**0,003871**	**0,003871**	**0,003871**	**0,003871**	**0,003871**	**0,003871**

Tabela 10

Błędy otrzymanych prognoz w przedziale ich weryfikacji metoda LEM – porównanie dla poszczególnych wartości wymiaru zanurzenia

		$m = 2$	$m = 3$	$m = 4$	$m = 5$	MD
EUR	σ_T	**0,016567**	0,023091	0,027531	0,020713	0,019456
	I^2	**0,0000146945**	0,0000285463	0,0000405773	0,0000229679	0,0000202651
USD	σ_T	**0,03064**	0,037881	0,035752	0,060821	0,040174
	I^2	**0,0000817405**	0,000124946	0,000111294	0,000322088	0,000140525
JPY	σ_T	0,032819	**0,022733**	0,033566	0,069382	0,057354
	I^2	0,000106486	**0,0000510923**	0,000111386	0,000475908	0,000137945
GBP	σ_T	**0,031965**	0,06184	0,066743	0,093721	0,060324
	I^2	**0,0000263803**	0,0000987321	0,00011501	0,000226773	0,0000939506
AUD	σ_T	0,016732	0,020581	**0,016218**	0,02153	0,016832
	I^2	0,0000435733	0,0000659234	**0,0000409352**	0,0000725474	0,0000440933
DKK	σ_T	0,003096	0,002416	0,003977	0,003623	**0,002256**
	I^2	0,0000283639	0,0000172646	0,0000467964	0,0000388246	**0,0000150536**

Analiza wyników predykcji (tabele 4-10) wykazuje, że rząd dokładności wnioskowania w przyszłość był dobry. Świadczą o tym małe wartości średniej arytmetycznej popełnionych błędów prognozy oraz niskie wartości odchyleń standardowych błędów prognozy. Najmniejsze błędy prognozy d_T, D_T uzyskano dla wymiaru zanurzenia równego 2 dla walut GBP oraz USD. Dla JPY najdokładniejsze prognozy otrzymano dla parametru $d = 3$ lub $d = 4$. Dla DKK najtrafniejsze prognozy uzyskiwano najczęściej dla $d = 3$. Natomiast dla EUR najlepsze wyniki otrzymano dla wymiaru zanurzenia równego 2 lub 5, a także dla metody MD.

W tabeli 10 przedstawiono średni błąd prognozy oraz współczynnik Theila wyznaczonych prognoz w całym przedziale weryfikacji. Pogrubioną czcionką zaznaczono wyniki, dla których błędy prognoz są najmniejsze.

Dla każdej wartości parametru d średnie błędy prognozy σ_T w całym przedziale weryfikacji są małe. Jednak najlepsze wyniki otrzymano dla wymiaru zanurzenia równego 2 (EURO, USD, GBP), 3 (JPY), 4 (AUD) lub MD (DKK). Współczynniki rozbieżności również są zbliżone do zera. Zatem można wnioskować, że otrzymane prognozy są dość dokładne. Analizując dane zawarte w tabeli 10, można również stwierdzić, że metoda MD daje dokładniejsze prognozy niż w przypadku gdy $d = 4$ lub 5.

Podsumowanie

W opracowaniu wyznaczono przyszłe wartości szeregów czasowych, utworzonych z wybranych kursów walut (AUD, DKK, EUR, GBP, JPY, USD), w okresie od 1.01.1999 roku do 30.10.2004 roku obejmującym 1474 notowania. Celem artykułu było porównanie efektywności wybranych metod predykcji pod względem dokładności wyznaczonych prognoz. Do prognozowania przyszłych notowań wykorzystano metodę „najbliższych sąsiadów" oraz metodę LEM. W pracy porównano również wyniki wnioskowania w przyszłość metodą LEM dla różnych wartości wymiaru zanurzenia.

Literatura

1. Casdagli M.: Chaos and Deterministic versus Stochastic Non-linear Modelling. "J. R. Statist. Soc.54" 1991, No. 2, s. 303-328.
2. Chun S.H., Kim K.J., Kim S.H.: Chaotic Analysis of Predictability versus Knowledge Discovery Techniques: Case Study of the Polish Stock Market. "Expert Systems" 2002, Vol. 19, No. 5, s. 264-272.
3. Farmer J.D., Sidorowich J.J.: Exploiting Chaos to Predict the Future and Reduce Noise. In: Evolution, Learning and Cognition. Ed. Y.C. Lee. World Scientific, Singapore 1988.
4. Farmer J.D., Sidorowich J.J.: Predicting Chaotic Time Series. "Physical Review Letters" 1987, Vol. 59, No. 8, s. 845-848.
5. Miśkiewicz M.: Wrażliwość na zmianę warunków początkowych w szeregach czasowych kursów akcji notowanych na GPW w Warszawie. W: Postępy Ekonometrii. Red. A.S. Barczak. AE, Katowice 2004, s. 193-200.
6. Peters E.E.: Teoria chaosu a rynki kapitałowe. Nowe spojrzenie na cykle, ceny i ryzyko. WIG-Press, Warszawa 1997.
7. Takens F.: Detecting Strange Attractors Turbulence. Springer-Verlag, Berlin 1981.
8. Wolf A., Swift J.B., Swinney H.L., Vastano J.A.: Determining Lyapunov Exponents from a Time Series. „Physica 16D" 1985, s. 285-317.
9. Welfe A.: Ekonometria. Metody i ich zastosowanie. Polskie Wydawnictwo Ekonomiczne, Warszawa 2003.
10. Weron A., Weron R.: Inżynieria finansowa. Wydawnictwa Naukowo-Techniczne, Warszawa 1998.
11. Zawadzki H.: Chaotyczne systemy dynamiczne. Elementy teorii i wybrane przykłady ekonomiczne. AE, Katowice 1996.

12. Zeliaś A.: Teoria prognozy. Polskie Wydawnictwo Ekonomiczne, Warszawa 1997.
13. Zeug K.: Rekonstrukcja przestrzeni stanów na podstawie jednowymiarowego ekonomicznego szeregu czasowego. „Studia Ekonomiczne" 2005, nr 36, s. 227-241.
14. Zhang J., Lam K.C., Yan W.J., Gao H., Li Y.: Time Series Prediction Using Lyapunov Exponents in Embedding Phase Space. „Computers and Electrical Engineering 30" 2004, pp. 1-15.

THE „NEAREST NEIGHBOURS" METHOD VERSUS „LEM METHOD" – COMPARISON OF THE EFFECTIVENESS OF THE METHODS PREDICTING THE ECONOMIC PHENOMENA DESCRIBED BY TIME SERIES

Summary

This paper describes two methods of time series prediction. First method uses composite neighbours. A predicted value is defined by some weighted combination of the actual nearest neighbours in the reconstructed state space. Second method is based on the fundamental characteristic behavior that a sensitive dependence upon initial conditions (SDUIC) and Lyapunov exponents are measure of the SDUIC in chaotic systems. This is done firstly by reconstructing a phase space using time series, then using Lyapunov exponents as qualitative parameter to predict an unknown phase space point. After transferring the phase space point, the predicted time series data can be obtained.The numerical example has shown that the first method is effective.

Anna Mularczyk

FRAKTALNE WSPOMAGANIE ZARZĄDZANIA ZAPASAMI

Wstęp

Zapasy, czyli „[...] dobra znajdujące się w dyspozycji przedsiębiorstwa, przeznaczone (lecz czasowo nie włączone) do działalności produkcyjnej" [1, s. 181] są nieodłączną częścią aktywności przedsiębiorstwa. Z zapasami, a raczej polityką ich zarządzania wiąże się zatem ściśle ryzyko. W niniejszym artykule ryzyko jest rozumiane jako możliwość odchyleń wyniku od stanu oczekiwanego będąca skutkiem zarówno konkretnego postępowania ludzkiego (motywowanego chęcią zysku bądź zniwelowania prawdopodobnej straty), jak i działania niezależnych czynników losowych. Wobec powyższego istnienie ryzyka pociąga za sobą koszty. W gospodarce materiałowej koszty magazynowania są przeciwwagą kosztów niedoboru zapasów, czyli kar za niezrealizowane zamówienie, bądź kosztów utraty klientów na rzecz konkurencji itp. Dąży się do tego, aby koszty w sumie były jak najmniejsze, aby zapewnić przedsiębiorstwu maksymalny zysk. W warunkach idealnych nie istnieje konieczność magazynowania zapasów, ponieważ są one dostarczane w odpowiednim czasie i w odpowiednich ilościach tak, by być od razu wykorzystywane w procesie produkcyjnym, realizując w trybie natychmiastowym zamówienia klientów. Jednak sytuacja taka jest najczęściej niemożliwa do osiągnięcia, ze względu na czas realizacji zamówień i dostawy materiałów oraz typ produkcji. Magazynowanie materiałów jest zatem koniecznością. Dlatego ważne jest znalezienie takiej wielkości zapasów, która zapewnia maksymalne korzyści, przy minimalnych nakładach, wiążąc się jednocześnie z minimalnym ryzykiem. Dodatkowym czynnikiem utrudniającym proces planowania – a więc zwiększającym ryzyko złego oszacowania zapotrzebowania – jest coraz większa zmienność otoczenia. Rosnąca turbulencja otoczenia jest wywołana m.in. globalizacją związaną z postępem

technicznym oraz zmianami politycznymi, co pociąga za sobą fakt, iż przed współczesnymi przedsiębiorstwami są stawiane bardzo wysokie wymagania. Z tego względu, aby utrzymać swoją pozycję na rynku, powinny one nieustannie analizować podejmowane decyzje – także te dotyczące gospodarki materiałowej – pod względem towarzyszącego im ryzyka. Oszacowanie i ocena ryzyka związanego z gospodarką materiałową może zatem pozwolić na odpowiedni dobór metody zarządzania zapasami.

Należy tu zwrócić uwagę na fakt, iż większość metod zarządzania zapasami*, uznając, iż zapotrzebowanie na nie jest losowe, przyjmuje założenie o zadanym jego rozkładzie, którym najczęściej jest rozkład normalny.

W artykule przedstawiono metodę analizy i oceny ryzyka na podstawie miary, jaką jest wymiar fraktalny. Metodę tę nazwano analizą ARRS. Istotne jest, iż nie ma w niej konieczności przyjmowania założenia o normalności rozkładu badanego szeregu. Dzięki temu może ona wnieść dodatkowe informacje o badanym zjawisku, bez konieczności przyjmowania upraszczających założeń. Metoda ta wywodzi się z teorii chaosu (co ma duże znaczenie ze względu na burzliwość otoczenia współczesnych przedsiębiorstw), a ściślej rzecz biorąc – z geometrii fraktalnej.

W dalszej części artykułu zostaną scharakteryzowane fraktale oraz wymiar fraktalny, tak aby kolejno móc je połączyć z pojęciem ryzyka towarzyszącego działalności przedsiębiorstwa i przedstawić metodę ARRS. Następnie zostaną przedstawione wyniki badań empirycznych przeprowadzonych w pewnym przedsiębiorstwie produkcyjnym oraz wyciągnięte wnioski.

1. Fraktale i wymiar fraktalny

Fraktale są obiektami geometrycznymi pojmowanymi zazwyczaj intuicyjnie. Benoit Mandelbrot uznawany za „ojca" geometrii fraktalnej odżegnuje się od podania konkretnej definicji fraktali [6, s. 18]. W literaturze przedmiotu jednak wciąż podejmuje się różne próby definiowania tych obiektów – bardziej lub mniej precyzyjne, w zależności od potrzeb. Przykładowo, w „Nowej encyklopedii powszechnej" PWN fraktal został określony jako: „[...] skomplikowana figura geometryczna, o której na pierwszy rzut oka trudno powiedzieć, czy jest krzywą, powierzchnią, czy ma jeszcze większy wymiar; charakteryzuje ją swo-

* Opartych na modelach zarządzania zapasami, które dzieli się ogólnie na: modele deterministyczne (charakteryzują się znanym zapotrzebowaniem na poszczególne materiały) oraz modele probabilistyczne (które są modelami bardziej zbliżonymi do rzeczywistości – i stosowanymi szeroko w praktyce – nie występuje w nich bowiem założenie o znajomości popytu na produkty, a więc i zapotrzebowania na materiały; wartość ta jest zmienną losową o określonym rozkładzie i może być tylko oszacowana z określonym prawdopodobieństwem, na podstawie popytu w okresach poprzednich).

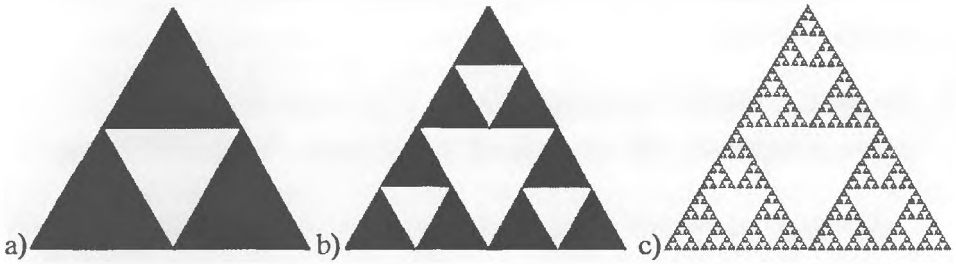
ista regularność w nieregularności – stopień tej regularności jest określony liczbą niecałkowitą (wymiar fraktalny)" [19]. J. Kurdewicz [4, s. 18] natomiast podaje, iż „[...] fraktalem na płaszczyźnie R^2 nazywamy każdy niepusty i zwarty podzbiór płaszczyzny R^2". Następnie zaznacza, iż można uogólnić tę definicję do przestrzeni metrycznej zupełnej X. Z kolei E. Peters [14, s. 7] twierdzi, iż: „[...] fraktal jest obiektem, którego części pozostają w pewnej relacji do całości".

W każdej z powyższych definicji została zaakcentowana jakaś – istotna dla jej autora – cecha fraktali.

Można stwierdzić, iż większość typowych fraktali charakteryzuje się trzema cechami:

1) samopodobieństwem – część figury przypomina całość w pewnej skali,
2) prostotą konstrukcji – są tworzone za pomocą wzorów rekurencyjnych,
3) niecałkowitym wymiarem (a przynajmniej większym od ich wymiaru w sensie topologicznym*).

Na poniższym rysunku zaprezentowano przykładowy fraktal, jakim jest trójkąt Sierpińskiego (rys. 1). Rysunki a i b przedstawiają pierwsze kroki jego konstrukcji, natomiast na rysunku c przedstawiono jego obraz w nieskończoności (oczywiście uwzględniając rozdzielczość możliwą w druku), czyli wspomniany fraktal.



Rys. 1. Trójkąt Sierpińskiego (c) i sposób jego konstrukcji (a) i (b)

Źródło: [18].

Najbardziej istotną cechą fraktali – z punktu widzenia analizy i oceny ryzyka – jest wymiar fraktalny. Jak już wspomniano, jest on zazwyczaj liczbą ułamkową – jego wartość dostarcza informacji o tym, w jaki sposób dany obiekt wypełnia przestrzeń, w której jest zanurzony. Im wymiar fraktalny jest bliższy wymiarowi tej przestrzeni, w tym większym stopniu ją wypełnia**.

---

* Istnieją bowiem fraktale o całkowitym wymiarze, należą do nich tzw. diabelskie schody, krzywa Hilberta czy brzeg zbioru Mandelbrota – ich wymiar fraktalny wynosi 2, jednakże ich wymiary topologiczne są równe 1. Por. np. [12].

** Przedstawiony na rys. 1 Trójkąt Sierpińskiego o wymiarze topologicznym 1 znajduje się w 2-wymiarowej przestrzeni, a jego wymiar fraktalny wynosi 1,585.

Klasycznym przykładem zaczerpniętym z literatury [5; 6; 12; 18] jest pomiar linii brzegowej Wielkiej Brytanii. Dokonując tego pomiaru za pomocą miar (odcinków, ramion cyrkla) o różnych długościach otrzymuje się różne wyniki. Im mniejsza użyta miarka jednostkowa (mniejszy rozstaw cyrkla), tym większy wynik, co jest skutkiem zwiększanej dokładności pomiaru, ponieważ coraz mniejsze elementy (od zatok, po głazy, kamienie, ziarnka piasku itd.) są brane od uwagę. Nie jest możliwe zatem podanie dokładnej długości linii brzegowej czy naturalnej granicy jakiegokolwiek państwa (gdyż zawsze jest możliwe użycie mniejszej miarki), ponieważ w rzeczywistości nie jest ona zbieżna do żadnej wartości, lecz dąży do nieskończoności. Jednak, jak zauważył Mandelbrot, każda z tych granic ma pewną sobie charakterystyczną cechę, za pomocą której można ją opisać. Jest to pewna stała zależność, w jakiej pozostają dana długość miarki i wyznaczona długość granicy, czyli właśnie wymiar fraktalny*:

$$N_\varepsilon = c\,\varepsilon^{-D} \tag{1}$$

gdzie:
$N\varepsilon$ – liczba odcinków,
c – stała,
ε – długość odcinka (miarki, rozstaw ramion cyrkla),
D – wymiar fraktalny.

2. Ocena ryzyka związanego z gospodarką materiałową za pomocą wymiaru fraktalnego

Narzędzia geometrii fraktalnej zostały już niejednokrotnie zastosowane w ekonomii. W literaturze przedmiotu można znaleźć liczne tego przykłady**. Aplikacja metod fraktalnych do analizy finansowych szeregów czasowych na giełdach, zarówno na rynkach zagranicznych, jak i polskich, daje wiele nowych możliwości. W literaturze przedmiotu opisano również aplikację narzędzi geometrii fraktalnej do organizacji, czego przykładem jest koncepcja organizacji fraktalnej obszernie opisana w pracach [2; 3;13;16].

Głównym narzędziem wykorzystywanym podczas analiz rynków finansowych jest wymiar fraktalny, który ze względu na swe właściwości może także służyć do oceny wielkości ryzyka.

* Wymiar fraktalny wybrzeża Wielkiej Brytanii wynosi około 1,319.
** Np. w [7; 14]. Szerokiego przeglądu literatury z zakresu powyższego tematu dokonali autorzy opracowań [9] oraz [10].

Analogicznie, badając ryzyko w przedsiębiorstwie, również można do tego celu wykorzystać pewne własności wymiaru fraktalnego. Wielkości zapotrzebowania na materiały zmieniają się wraz z upływem czasu, tworząc szeregi czasowe. Szeregi te po naniesieniu na układ współrzędnych są liniami łamanymi, czyli obiektami geometrycznymi, których wymiar topologiczny wynosi 1. Ich wymiar fraktalny jest natomiast liczbą z przedziału (1; 2). Im zatem bliższy jest on 1, tym mniej zygzakowaty jest dany wykres (zbliżony do linii prostej). Taki szereg czasowy jest łatwy do predykcji – można z niego łatwo odczytać trend i przewidzieć przyszłe zmiany. Natomiast wymiar fraktalny bliższy 2 – odwrotnie – oznacza, że dany szereg jest bardziej „postrzępiony", czyli w większym stopniu wypełnia obszar wykresu. W takiej sytuacji istnieje wyższe prawdopodobieństwo częstszych zmian badanego zjawiska. Szereg taki jest szeregiem chaotycznym, a więc mniej przewidywalnym. Zatem, analizując szeregi reprezentujące wielkość zapotrzebowania na dany materiał, należy wyznaczyć ich wymiary fraktalne – wówczas: wyższy wymiar oznacza większe ryzyko związane z badanym materiałem. Wymiar fraktalny funkcjonuje tu jako detektor sygnałów płynących zarówno z otoczenia, jak i samego procesu produkcyjnego. Sygnały te mają różne nasilenie ze względu na chaotyczny charakter otoczenia oraz zakłóceń towarzyszących nieustannie procesowi produkcyjnemu*.

Jednym ze sposobów wyznaczania wymiaru fraktalnego szeregów czasowych jest analiza przeskalowanych zakresów zwana inaczej analizą R/S**. Analiza R/S jest fundamentalną częścią przedstawianej metody analizy ryzyka (ARRS), dlatego właśnie za jej pomocą szacuje się wymiar fraktalny. W wyniku jej zastosowania otrzymuje się przede wszystkim wykładnik Hursta (*H*), który oprócz wskazania na wartość szukanego wymiaru fraktalnego wnosi niezwykle istotne informacje o badanym szeregu.

Analiza ARRS przebiega w sześciu krokach:

KROK 1

Zebranie danych*** dotyczących rozchodów z magazynu surowców w postaci szeregów czasowych.

---

* Podobną sytuację, w zastosowaniu do analizy rynków kapitałowych, opisał E. Peters [14, s.62]: „Sposób wypełnienia przestrzeni przez dany obiekt zależy od sił biorących udział w jego kształtowaniu. [...] W przypadku szeregów czasowych stóp zwrotu akcji siłami tymi będą dane mikro- i makroekonomiczne, które wpływają na sposób, w jaki inwestorzy postrzegają wartość aktywów".

** Analiza ta została opracowana przez angielskiego hydrologa zajmującego się pływami Nilu, H.E. Hursta, w połowie XX wieku. Stosując ją, udowodnił, że pływy Nilu nie tylko nie są niezależne od siebie, ale występują w nich długookresowe trendy [12; 14].

*** Liczba obserwacji powinna być duża (co najmniej kilka setek) – zgodnie z wymaganiami analizy R/S.

KROK 2

Zamiana danych na wahania*, zgodnie z równaniem (2):

$$z_t = \ln\left(\frac{x_t}{x_{t-1}}\right) \tag{2}$$

gdzie:

z_t – logarytmiczna stopa zwrotu,

x_t – wartość w chwili t.

KROK 3

Przeprowadzenie obliczeń związanych z analizą przeskalowanych zakresów (rys. 2).

KROK 4

Wyznaczenie wymiaru fraktalnego według zależności (3):

$$D = 2 - H \tag{3}$$

gdzie:

D – wymiar fraktalny,

H – wykładnik Hursta.

KROK 5

Wyznaczenie wartości oczekiwanej wykładnika Hursta według poprawki [7; 15; 17]:

1. Dla podszeregów o małej długości, czyli dla $n < 30$, przybliżenie wartości oczekiwanej przeskalowanego zakresu otrzymuje się z równania (4):

$$E(R/S_n) = \frac{n-\frac{1}{2}}{n} \cdot \frac{\Gamma\left(\frac{n-1}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{n}{2}\right)} \cdot \sum_{i=1}^{n-1} \sqrt{\frac{n-i}{i}} \tag{4}$$

gdzie $\Gamma(n)$ jest funkcją gamma Eulera.

2. Natomiast dla podszeregów o długości $n \geq 30$ – ze wzorów (5-7):

$$E(R/S_n) = \left(n-\frac{1}{2}\right) \cdot G2(n) \cdot S2(n) \tag{5}$$

* Gdyż to właśnie nieoczekiwane wahania są źródłem ryzyka.

gdzie:

$$G2(n) = \sqrt{\frac{2}{n\pi}} \cdot \frac{n - 0{,}2491}{n - 1} \tag{6}$$

$$S2(n) = \frac{\pi}{2} - \frac{2{,}4718 - 3{,}5466n + 1{,}4635n^2}{(n-1)^{2{,}5}} \tag{7}$$

KROK 6

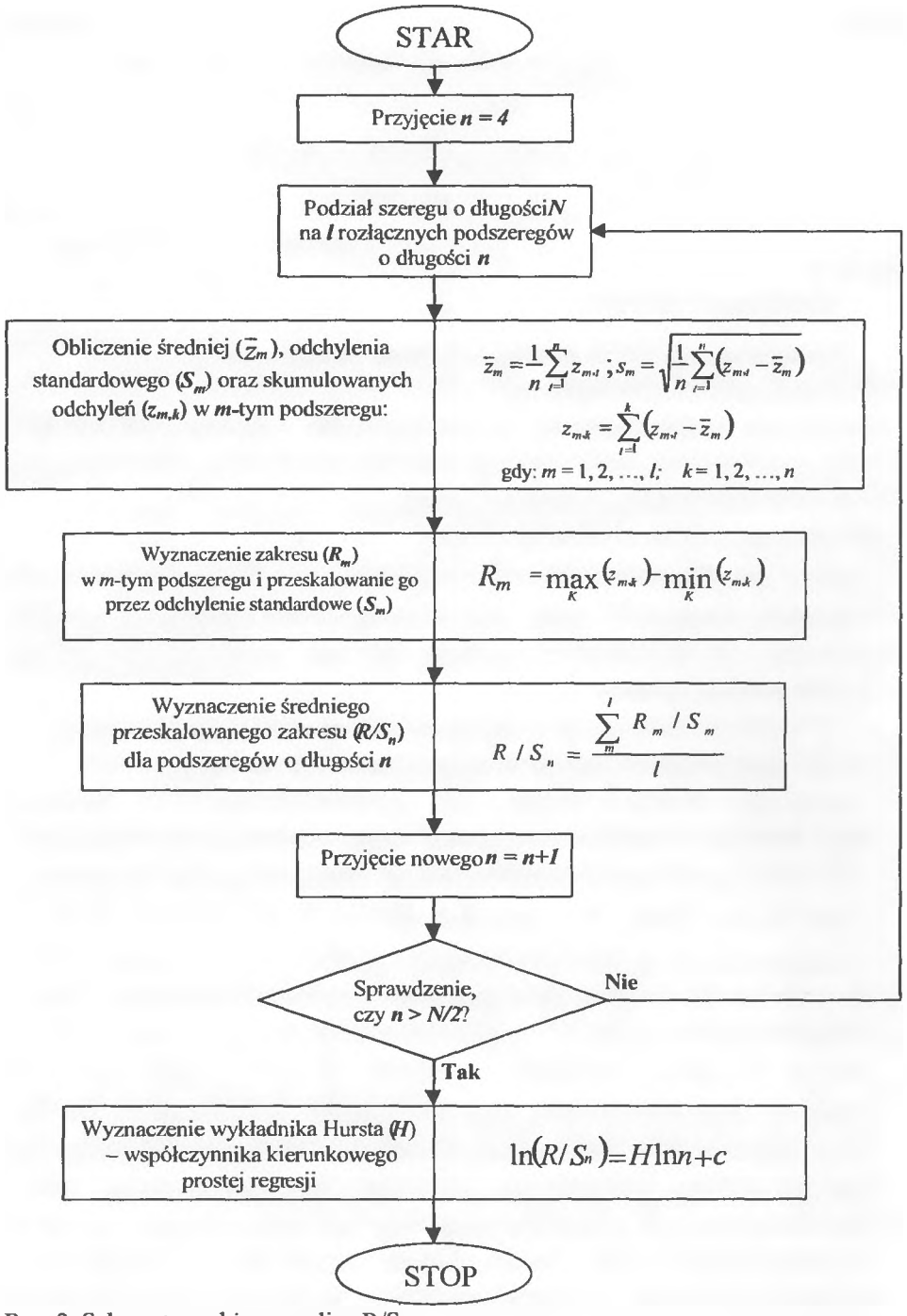
Interpretacja wyników.

Interpretacja wyników przebiega w dwóch płaszczyznach:

1. Za pomocą analizy porównawczej:
   Im większy wymiar fraktalny, tym większe ryzyko towarzyszy badanej dziedzinie. Porównując zatem wymiary fraktalne dwóch szeregów, można określić relatywnie bardziej ryzykowny obszar.
2. Interpretując wartość wykładnika Hursta:
   Ogólnie: gdy H jest różny od $E(H)$ co najmniej o $1/\sqrt{N}$ (gdzie N jest liczbą wszystkich obserwacji), wtedy badany szereg nie jest klasycznym procesem losowym – nie ma rozkładu normalnego. Wówczas można mieć do czynienia z jedną z dwóch sytuacji:
   - $H > E(H)$, co oznacza, iż w szeregu występuje efekt długiej pamięci,
   - $H < E(H)$, co jest wynikiem większej zmienności szeregu.

   Szczególnie: wykładnik Hursta może przyjmować wartości z przedziału [0, 1]. Pozwala on zaklasyfikować badany szereg do jednego z trzech rodzajów:

1. Gdy $H = 0{,}5$, czyli wymiar fraktalny szeregu wynosi 1,5 (zatem – dla szeregów o niezbyt dużej liczbie obserwacji: $H \in \left(E(H) - \left(1/\sqrt{N}\right), E(H) + \left(1/\sqrt{N}\right)\right)$, szereg jest szeregiem losowym, mającym charakter błądzenia przypadkowego, czyli ruchem Browna. Dane są zgodne z rozkładem normalnym. Istnieje jednakowe prawdopodobieństwo wzrostu i spadku wartości badanej cechy.
2. Gdy $H < 0{,}5$ (czyli odpowiednio: $H < E(H) - \left(1/\sqrt{N}\right)$), badany szereg jest szeregiem antypersystentnym, mającym własność „powracania do średniej". Im H bliższy jest 0, a zatem wymiar fraktalny bliższy 2, tym zachowanie szeregu jest bardziej chaotyczne niż w przypadku błądzenia losowego. Szeregi takie charakteryzują się bardzo postrzępioną linią, często dochodzi w nich do odwracania trendu, czyli – inaczej mówiąc – wyodrębnienie jakiegokolwiek trendu jest niemożliwe. Prawdopodobieństwo, że po wzroście nastąpi spadek (lub odwrotnie), jest wyższe od prawdopodobieństwa dwóch jednokierunkowych zmian.



Rys. 2. Schemat przebiegu analizy R/S

Źródło: Opracowanie własne na podstawie [7; 8; 14; 17].

3. Natomiast w ostatnim przypadku, wówczas gdy $H > 0{,}5$ (czyli: $H > E(H) + \left(1/\sqrt{N}\right)$), badany szereg jest persystentny, co oznacza, iż wzmacnia swój trend. Wymiar fraktalny zawiera się w przedziale (1,0; 1,5). W takiej sytuacji istnieje wyższe prawdopodobieństwo nastąpienia po sobie dwóch takich samych (co do kierunku) zmian niż przeciwnych. Szeregi persystentne są ułamkowymi ruchami Browna – przykładami obciążonego błądzenia przypadkowego z siłą obciążenia tym większą, im H bliższy 1.

Wartość wykładnika Hursta jest zatem prawdopodobieństwem – a raczej miarą obciążenia – z jakim nastąpią po sobie dwie jednakowe (jednokierunkowe) zmiany w badanym szeregu [14].

3. Analiza zużycia zapasów jako przykład analizy ARRS

Badaniom poddano zapotrzebowanie na materiały w pewnym przedsiębiorstwie. Zanalizowano zapotrzebowanie na walcówkę – będącą surowcem, z którego wytwarza się wszystkie wyroby badanego przedsiębiorstwa.

Należy w tym miejscu podkreślić wagę oraz duże ryzyko związane z odpowiednim gospodarowaniem zapasami walcówki, gdyż od tego zależy nie tylko powodzenie realizacji zamówień, ale i również fizyczna możliwość ich przyjęcia. Ponieważ wiele z zamówień kierowanych do przedsiębiorstwa bywa składanych w „trybie natychmiastowym", częstokroć – bez posiadania odpowiednich rezerw walcówki – przedsiębiorstwo nie ma odpowiedniego potencjału do ich przyjęcia. Pozyskanie walcówki wiąże się bowiem, oprócz okresu realizacji zamówienia i dostawy (co często jest związane z wyższymi kosztami wynikającymi m.in. z windowania cen surowca przez dostawców – ze względu na pilność zamówienia), z kilkudniowym okresem „leżakowania" surowca. W związku z tym brak odpowiedniej ilości zapasów walcówki jest obarczony ryzykiem utraty klienta – na rzecz konkurencji. Z kolei przetrzymywanie zbyt dużej ilości zapasów wiąże się z ryzykiem zbyt dużego zamrożenia środków płatniczych oraz ponoszeniem kosztów magazynowania.

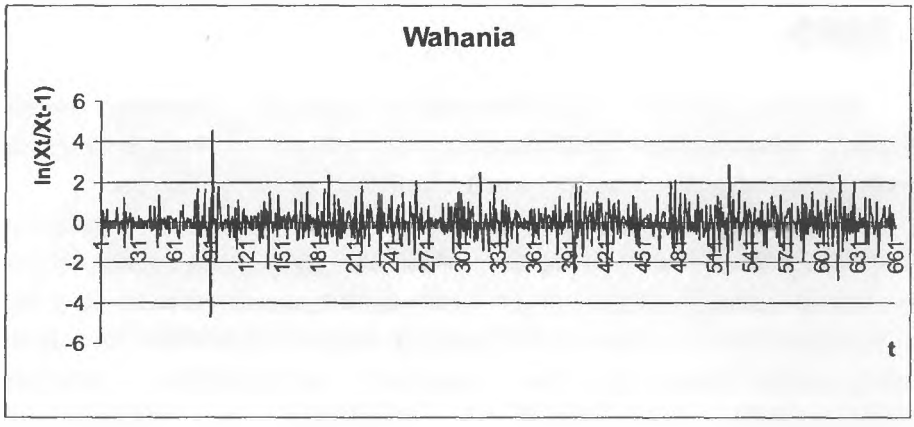
Walcówkę można podzielić – ze względu na typ realizowanej w przedsiębiorstwie produkcji – głównie na dwa rodzaje: walcówkę „cienką", której średnica nie przekracza 6,5 mm, oraz walcówkę „grubą" – o średnicach większych.

Zebrano dane dotyczące:

1. Rozchodów walcówki z magazynu ogółem, w okresie: 1 VIII 2003-25 V 2006; 662 dziennych obserwacji.
2. Rozchodów walcówki „cienkiej" z magazynu, w okresie: 1 VIII 2003-25 V 2006; 575 dziennych obserwacji.
3. Rozchodów walcówki „grubej" z magazynu, w okresie: 1 VIII 2003-25 V 2006; 576 dziennych obserwacji.

Walcówka ogółem

Wahania zmian poziomu zapotrzebowania na walcówkę ogółem przedstawiono na rys. 3.



Rys. 3. Wahania zmian

W pierwszej kolejności zbadano zgodność danego szeregu z rozkładem normalnym*. Szereg wykazał brak zgodności z rozkładem normalnym**, co

---

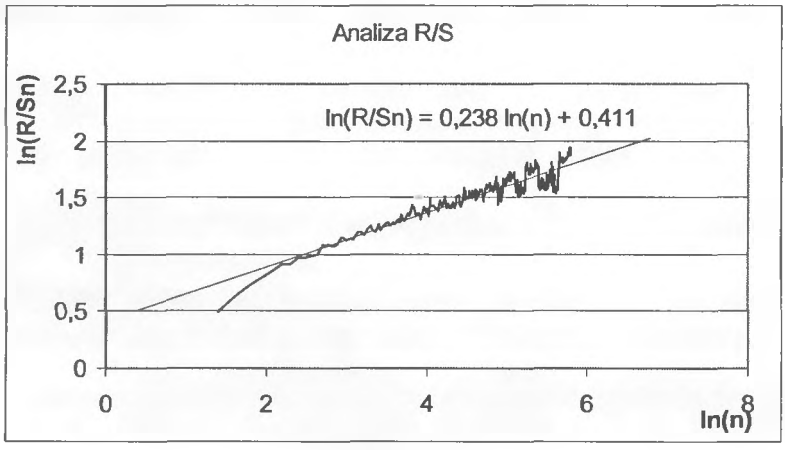
* W tym celu zastosowano procedurę nieparametrycznego testu zgodności *chi*-kwadrat (χ^2) [11]. Zgodnie z jego założeniami, weryfikowana hipoteza ma postać: H_0: *rozkład Z_t ma rozkład normalny.* Wobec hipotezy alternatywnej: H_1: *rozkład Z_t nie ma rozkładu normalnego.* Sprawdzianem hipotezy H_0 jest statystyka określona równaniem: $\chi^2 = \sum_{i=1}^{r} \frac{(n_i - np_i)^2}{np_i}$, gdzie χ^2 to wartość empiryczna statystyki, n_i – liczebność *i*-tego przedziału klasowego (liczebności empiryczne), np_i – liczba jednostek, które powinny znaleźć się w *i*-tym przedziale klasowym, przy założeniu, że cecha ma rozkład normalny (liczebności teoretyczne). Relacja wyznaczająca zbiór krytyczny ma postać: $P\left(\chi^2 > \chi_\alpha^2\right) = \alpha$, gdzie χ_α^2 jest wartością krytyczną $k = r - s - 1$ stopni swobody i $P = \alpha$, przy czym s oznacza liczbę parametrow, które należy wstępnie wyznaczyć, a r – liczbę przedziałów klasowych.

** W wyniku przeprowadzonego testu odrzucono hipotezę H_0 na rzecz hipotezy alternatywnej (H_1), gdyż $\chi^2 = 16{,}85$ wobec $\chi_\alpha^2 = 15{,}51$ (dla $\alpha = 0{,}05$), czyli: $\chi^2 > \chi_\alpha^2$.

oznacza, iż w takim przypadku wysoce pomocne powinno być zastosowanie analizy fraktalnej, czyli metody ARRS.

Wynik analizy przeskalowanych zakresów jest widoczny na wykresie* (rys. 4).

Wykładnik Hursta badanego szeregu wynosi $H = 0{,}238$, zatem wymiar fraktalny badanego szeregu zgodnie z równaniem (3) wynosi $D = 1{,}762$.



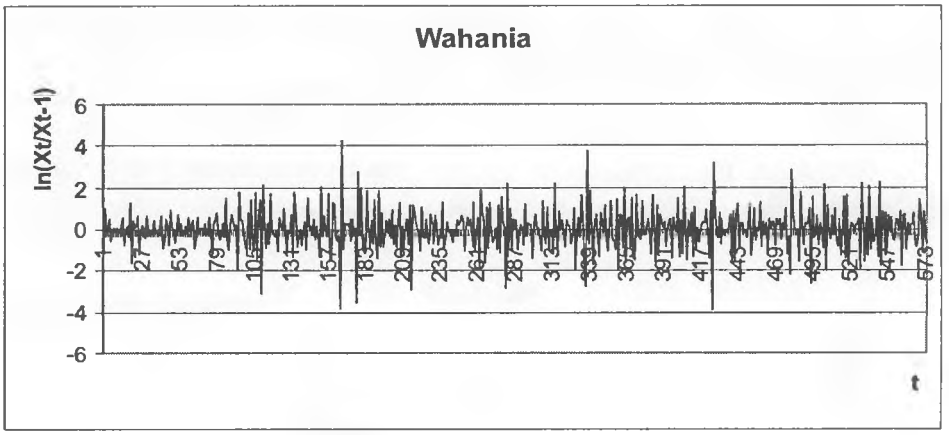
Rys. 4. Analiza R/S rozchodów walcówki ogółem

Po dodatkowych obliczeniach (równania 4-7) otrzymano wartość oczekiwaną wykładnika Hursta: $E(H) = 0{,}545$ oraz wartość krytyczną: $1/\sqrt{N} = 0{,}039$. Wykładnik Hursta znajduje się poza przedziałem krytycznym, który wynosi (0,505; 0,584), co oznacza, iż szereg nie jest błądzeniem losowym – ma on charakter bardziej chaotyczny i mniej przewidywalny, czyli jest obarczony większym ryzykiem.

Walcówka „cienka"

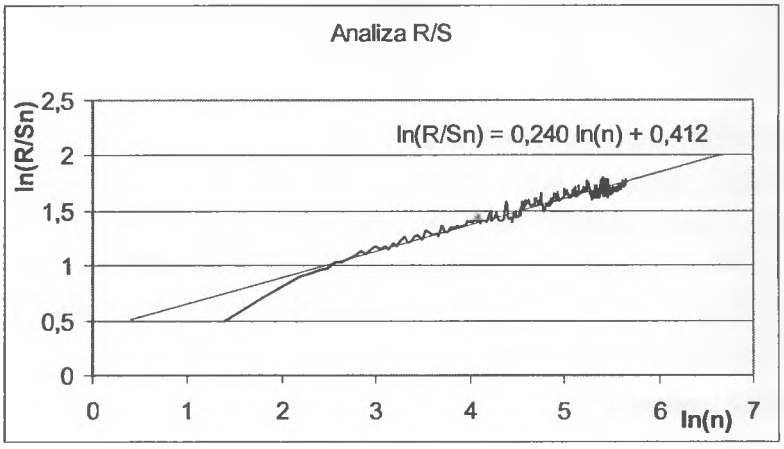
W dalszej kolejności badaniom poddano rozchody z magazynu walcówki oznaczonej jako „cienka". Wahania poziomów zapotrzebowania na ten surowiec przedstawiono na rys. 5.

* Wykres jest wykresem podwójnie logarytmicznym obrazującym zależność $\ln(R/S_n)$ od $\ln(n)$. Na punkty empiryczne została na nim naniesiona prosta regresji.



Rys. 5. Wahania zmian poziomu zapotrzebowania na walcówkę „cienką"

Szereg wielkości zapotrzebowania na walcówkę „cienką" również nie jest zgodny z rozkładem normalnym*. Wyniki analizy R/S zostały zaprezentowane na rys. 6.



Rys. 6. Analiza R/S rozchodów walcówki „cienkiej"

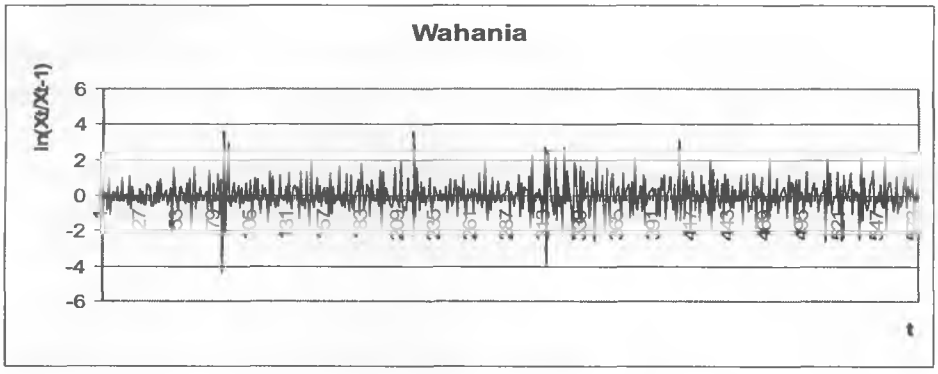
Wynik ten jest bardzo zbliżony do wyniku analizy walcówki ogółem: wykładnik Hursta wyniósł $H = 0{,}240$, wobec czego wymiar fraktalny badanego szeregu: $D = 1{,}760$.

* $\chi^2 = 39{,}58$ wobec $\chi_\alpha^2 = 18{,}31$ (dla $\alpha = 0{,}05$), czyli: $\chi^2 > \chi_\alpha^2$.

Po dodatkowych obliczeniach (równania 4-7) otrzymano wartość oczekiwaną wykładnika Hursta na poziomie: $E(H)$ = 0,546 oraz wartość krytyczną: $1/\sqrt{N}$ = 0,042, co pozwoliło na określenie przedziału krytycznego (0,504; 0,586). Zatem wykładnik Hursta znajduje się poza tym przedziałem, co oznacza, jak poprzednio, iż szereg nie jest błądzeniem losowym. Badany szereg ma charakter mniej przewidywalny, czyli jest obarczony wysokim ryzykiem.

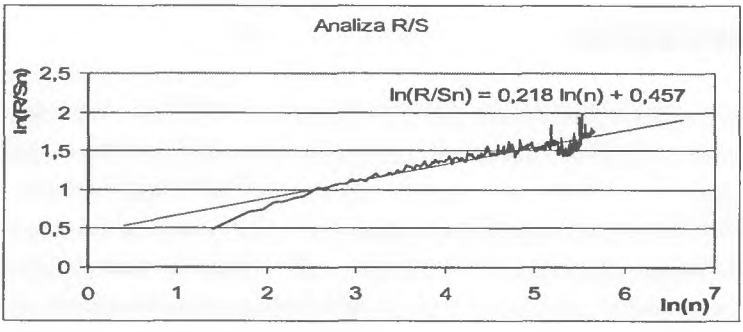
Walcówka „gruba"

W ostatniej części badaniom poddano rozchody z magazynu walcówki „grubej". Wykres wahań tych rozchodów znajduje się na rys. 7. Po weryfikacji odrzucono również hipotezę o normalności rozkładu badanego szeregu*.



Rys. 7. Wahania zmian poziomu zapotrzebowania na walcówkę „grubą"

Wykres obrazujący analizę przeskalowanych zakresów znajduje się na rys. 8.



Rys. 8. Analiza R/S rozchodów walcówki „grubej"

* $\chi^2 = 17{,}13$ wobec $\chi_\alpha^2 = 16{,}92$ (dla $\alpha = 0{,}05$), czyli: $\chi^2 > \chi_\alpha^2$.

Wykładnik Hursta badanego szeregu wynosi $H = \mathbf{0,218}$, zatem wymiar fraktalny: $D = \mathbf{1,782}$, czyli jest nieznacznie większy od wymiarów fraktalnych szeregów poprzednich.

Ponadto, po uwzględnieniu poprawki, wartość oczekiwana wykładnika Hursta, jak w przypadku walcówki „cienkiej" wynosi $E(H) = \mathbf{0,546}$, a wartość krytyczna: $1/\sqrt{N} = 0,042$. Zatem wykładnik Hursta znajduje się poza przedziałem krytycznym, który wynosi (0,504; 0,586), co oznacza, iż szereg tak samo jak w poprzednim przypadku nie jest błądzeniem losowym i wykazuje większą zmienność.

Wnioski

Wszystkie przebadane szeregi nie są zgodne z rozkładem normalnym oraz wykazują dużą zmienność. Zmienność ta jest ponadto wyższa niż w przypadku błądzenia losowego. Istnieje zatem znaczne ryzyko nieoczekiwanych zmian, które mogłyby zakłócić gospodarkę zapasami materiałowymi. Dodatkowo (porównując odpowiednie wymiary fraktalne) wiadomo, iż wyższy poziom ryzyka jest związany z zapotrzebowaniem na walcówkę „grubą", a nie „cienką", co powinno być brane pod uwagę podczas planowania wielkości zapasów. Również faktem jest wykrycie zjawiska chaosu w badanej sferze (wszystkie szeregi mają charakter chaotyczny). Może to być odzwierciedleniem burzliwości rynku zaopatrzenia lub rynku zbytu przedsiębiorstwa i jest związane z dużym ryzykiem. Wobec powyższego zalecane jest głębsze przyjrzenie się obu rodzajom badanego surowca i monitoring zmian wielkości związanych z gospodarowaniem badaną walcówką (szczególnie drugiego typu) w celu minimalizacji ryzyka.

Podsumowanie

Burzliwość i chaotyczność współczesnego rynku jest główną przesłanką przemawiającą za wykorzystaniem metod wywodzących się z teorii chaosu, jakimi są metody fraktalne w przedsiębiorstwach. Wymiar fraktalny jest miarą bardzo wrażliwą na niejednorodność badanego zjawiska – pozwala on na detekcję zarówno długoterminowych zależności, jak i chaosu w szeregu, którego wartości są kształtowane zarówno przez bliższe (samo przedsiębiorstwo), jak i dalsze otoczenie (otoczenie przedsiębiorstwa), a więc może być wykorzystywany jako miara ryzyka.

W niniejszym artykule przedstawiono zastosowanie fraktalnej analizy ryzyka (ARRS) w gospodarce materiałowej. Jednakże metoda ta może być stoso-

wana w każdej mierzalnej sferze działalności przedsiębiorstwa. Dzięki znajomości wymiaru fraktalnego badanej dziedziny dany podmiot, dysponując wiedzą o charakterze zjawiska, jest w stanie adekwatnie reagować na zmiany zachodzące w otoczeniu. Następstwem tego jest uniknięcie stagnacji oraz zwiększenie efektywności.

Należy jednak zaznaczyć, iż analizując ryzyko metodami fraktalnymi, nie można zupełnie odrzucić wcześniej stosowanych i powszechnie uznanych metod. W niektórych przypadkach (np. braku normalności badanego szeregu) analiza ARRS może dać lepsze efekty, natomiast w pozostałych może stanowić czynnik komplementarny.

Literatura

1. Badania operacyjne. Red. E. Ignasiak. PWE, Warszawa 1996.
2. Jakimowicz A.: Organizacje fraktalne i chaotyczne. W: Multimedia w biznesie. Red. L. Kiełtyka. Zakamycze, Kraków 2003, s. 125-138.
3. Krawczyk A.: Geometria fraktalna jako instrument opisu organizacji. „Przegląd Organizacji" 2001, nr 4, s. 17-20.
4. Kurdewicz J.: Fraktale i chaos. Wydawnictwa Naukowo-Techniczne, Warszawa 1993.
5. Mandelbrot B.B.: The Fractal Geometry of Nature. W.H. Freedman and Company, New York, Berlin, Heidelberg 1982.
6. Mandelbrot B.B.: Gaussian Self-Afinity and Fractals: Globality, The Earth, 1/f Noise, and R/S. Springer-Verlag, New York, Berlin, Heidelberg 2002.
7. Mastalerz-Kodzis A.: Modelowanie procesów na rynku kapitałowym za pomocą multifraktali. AE, Katowice 2003.
8. Mularczyk A.: Analiza ryzyka w działalności przedsiębiorstwa z zastosowaniem metod fraktalnych. Politechnika Śląska, 2005.
9. Mulligan R.F.: Fractal Analysis of Highly Volatile Markets: an Application to Technology Equities. „The Quarterly Review of Economics and Finance" 2004, No. 44, pp. 155-179.
10. Mulligan R.F., Lobardo G A.: Maritime Businesses: Volatile Stock Prices and Market Valuation Inefficiencies. „The Quarterly Review of Economics and Finance" 2004, No. 44, pp. 321-336.
11. Ostasiewicz S., Rusnak Z., Siedlecka U.: Statystyka. Elementy teorii i zadania. Wydawnictwo AE im. Oskara Langego we Wrocławiu, Wrocław 1999.
12. Peitgen H.-O.: Jürgens H., Saupe D.: Fraktale. Granice chaosu, PWN, Warszawa 2002.

13. Perechuda K.: Organizacja fraktalna. W: Zarządzanie przedsiębiorstwem przyszłości. Koncepcje, modele, metody. Red. K. Perechuda. Placet, Warszawa 2000, s. 25-35.
14. Peters E.E.: Teoria chaosu a rynki kapitałowe. WIG-Press, Warszawa 1997.
15. Purczyński J.: Wybrane problemy numeryczne stosowania analizy R/S. „Przegląd Statystyczny" 2000, R. XLVII, Zeszyt 1-2, s. 17-21.
16. Warnecke H.-J.: Rewolucja kultury przedsiębiorstwa. Przedsiębiorstwo fraktalne. PWN, Warszawa 1999.
17. Weron A., Weron R.: Inżynieria finansowa. Wydawnictwa Naukowo-Techniczne, Warszawa 1998.
18. www.alpha.mini.pw.edu.pl/MiNIwyklady/fraktale (06.2006).
19. www.encyklopedia.pwn.pl/0_3.html (03.2005).

THE FRACTIONAL SUPPORT OF RESERVES MANAGEMENT

Summary

In the paper the fractional method of risk analysis (ARRS) as the support of reserves management has been presented. Turbulence of the contemporary market is the reason for the application of the fractional methods (as derived from chaos theory) into enterprises. The main element of the ARRS analysis is the rescaled range analysis – via one the fractional dimension is assessed. The fractional dimension is very susceptible measure to heterogeneity. The results of the researches carried out in a productive enterprise as an exemplary case of the proposed method have been presented. The demand for blank, which is the main resource in the enterprise, was examined. The importance and high risk concerned with the management of blank resources have been stressed. All of the examined sets of data had chaotic character and variation higher than in random walk which indicated high risk level.

Ewa Pośpiech

ZASTOSOWANIA ALGORYTMU KODUJĄCEGO RSA W BANKOWOŚCI ELEKTRONICZNEJ

Wstęp

Coraz większe możliwości, jakie daje korzystanie z powszechnie dostępnego sprzętu i techniki komputerowej, sprawiają, że przekazywanie informacji, zakupy internetowe czy korzystanie z bankowości elektronicznej stają się codziennością zarówno dla przedstawicieli różnych instytucji, jak i dla użytkowników prywatnych. Istotną rzeczą jest bezpieczeństwo przesyłanych informacji oraz dokonywanych transakcji. Ogromne zastosowanie w tej dziedzinie ma współczesna kryptografia, której rozwój jest możliwy dzięki zaawansowanym narzędziom matematycznym i informatycznym.

Celem artykułu jest przedstawienie jednego ze stosowanych w kryptografii algorytmów – systemu RSA (jeden z tzw. systemów jawnego klucza), który jest oparty na elementarnej teorii liczb, a jego bezpieczeństwo zależy od trudności rozkładania dużych liczb całkowitych na czynniki pierwsze.

1. Definicje i twierdzenia

Przedstawienie każdego zagadnienia, zwłaszcza takiego, które wykorzystuje aparat matematyczny, wymaga podania istotnych określeń przybliżających wprowadzane pojęcia. W tym celu niezbędne jest przytoczenie najważniejszych definicji i twierdzeń umożliwiających implementację pojęć w praktyce.

Twierdzenie 1 (Algorytm dzielenia)

Niech $a, b \in \mathbf{Z}$, $b > 0$. Wtedy istnieje dokładnie jedna para liczb całkowitych q i r, takich że $a = bq + r$, gdzie $0 \leq r < |b|$.

Liczbę q nazywa się ilorazem, a liczbę r resztą z dzielenia a przez b.

Dalsza część rozważań będzie się koncentrować na resztach z dzielenia liczb całkowitych. Technika pracy z resztami jest zwana arytmetyką modularną.

W arytmetyce modularnej mamy dodatnią liczbę całkowitą n, zwaną modułem. Dowolne dwie liczby całkowite, których różnica jest całkowitą wielokrotnością modułu, są traktowane jako równe lub równoważne z modułem. Wprowadza się następującą definicję.

Definicja 1

Jeśli n jest dodatnią liczbą całkowitą oraz a, b są liczbami całkowitymi, wówczas mówimy, że a jest przystające do b modulo n i piszemy:

$$a \equiv b \pmod{n}$$

jeśli $a - b$ jest liczbą całkowitą będącą wielokrotnością liczby n, tzn. $n \mid (a - b)$.

Dla dowolnego n każda liczba całkowita a jest równoważna dokładnie jednej z liczb 0, 1, 2, ..., $n - 1$. Liczbę tę nazywa się resztą z liczby a modulo n, a zbiór reszt modulo n jest oznaczany jako $Z_n = \{0, 1, 2, \ldots, n - 1\}$.

Jeśli a jest nieujemną liczbą całkowitą, to jej reszta modulo n jest resztą z dzielenia a przez n.

Twierdzenie 2*

Niech a i n będą liczbami całkowitymi i $n > 1$.

NWD$(a, n) = 1$ w $\mathbf{Z}$ wtedy i tylko wtedy, gdy równanie $ax = 1$ ma rozwiązanie w zbiorze Z_n, czyli $ax \equiv 1 \pmod{n}$, $x \in Z_n$**.

Lemat 1

Niech $p, r, s, c \in \mathbf{Z}$ oraz p to liczba pierwsza.

Jeśli $p \nmid c$ oraz $rc \equiv sc \pmod{p}$, wówczas $r \equiv s \pmod{p}$.

---

* W poniższym twierdzeniu stosuje się następujące oznaczenia: NWD (a, n) – największy wspólny dzielnik liczb a i n, $\mathbf{Z}$ – zbiór liczb całkowitych.

** Równanie można rozwiązać stosując np. algorytm Euklidesa (zob. [6]).

Lemat 2 (Małe twierdzenie Fermata)

Jeśli p jest liczbą pierwszą, $a \in \mathbf{Z}$ i $p \nmid a$, to:

$$a^{p-1} \equiv 1 \pmod{p}$$

Załóżmy, że p i q są różnymi dodatnimi liczbami pierwszymi. Niech $n = pq$ i $k = (p-1)(q-1)$. Wybierzmy d takie, że liczby d i k są względnie pierwsze, czyli NWD(d, k) = 1. Wówczas, na mocy twierdzenia 2, równanie $dx = 1$ ma rozwiązanie w Z_k.

A zatem, równanie:

$$dx \equiv 1 \pmod{k} \tag{1}$$

ma rozwiązanie w Z_k, które będzie oznaczone przez e.

Twierdzenie 3

Niech p, q, n, k, e, d są liczbami określonymi jak wyżej. Wówczas

$$b^{ed} \equiv b \pmod{n}, \ \forall\, b \in \mathbf{Z}$$

Podane wyżej twierdzenia mają duże znaczenie przy określaniu istoty algorytmu RSA.

2. Opis systemu RSA

RSA jest jednym z najpopularniejszych algorytmów kryptografii asymetrycznej*. Kryptografia asymetryczna jest obecnie powszechnie stosowana do wymiany informacji poprzez kanały o niskiej poufności, jak np. Internet, jednak najważniejsze zastosowania tego rodzaju kryptografii to przede wszystkim szyfrowanie i podpisywanie cyfrowe. Czynności te wymagają istnienia dwóch kluczy – prywatnego i publicznego, przy czym odtworzenie klucza prywatnego na podstawie klucza publicznego musi być obliczeniowo trudne. W przypadku szyfrowania, klucz publiczny, który jest udostępniony każdemu chcącemu zaszyfrować wiadomość, jest używany do zakodowania informacji, natomiast klucz prywatny – do jej odczytania.

W celu wygenerowania klucza RSA losuje się dwie duże liczby pierwsze p i q (w praktyce liczby te powinny być kilkudziesięciocyfrowe, takie by np. $pq > 10^{150}$) oraz liczbę d względnie pierwszą z $k = (p-1)(q-1)$. Rozwiązując równanie (1), uzyskuje się liczbę e, stanowiącą wraz z liczbą n klucz publiczny.

* Kryptografia asymetryczna to rodzaj kryptografii, w którym używa się zestawów dwóch lub więcej powiązanych ze sobą kluczy, umożliwiających wykonywanie różnych czynności kryptograficznych. Szyfry asymetryczne opierają się na istnieniu pewnych trudnych do odwrócenia problemów (np. łatwo dokonać mnożenia, trudniej rozłożyć na czynniki pierwsze) (na podstawie: http://pl.wikipedia.org).

Wiadomość, która ma być przesłana, musi być najpierw zamieniona na formę numeryczną przez zastąpienie każdej litery lub spacji (ewentualnie innych znaków) liczbą dwucyfrową. Przykładowy sposób przyporządkowania podaje tabela 1.

Tabela 1

spacja	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
P	Q	R	S	T	U	V	W	X	Y	Z	.	,	(	:	)
16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31

Numeryczny odpowiednik wiadomości niech będzie oznaczony literą B.

Niech p, q, n, k, e, d będą liczbami spełniającymi założenia twierdzenia 3 oraz niech dodatkowo liczby p i q są takie, że $B < pq = n$.

Aby zakodować wiadomość B, należy posłużyć się funkcją w postaci:

$$f(B) = B^e \text{ MOD } n = C$$

czyli obliczyć resztę z liczby B^e dzielonej przez n. Liczba C jest zakodowaną formą B. Osoba, która otrzymuje C, może ją zdekodować, obliczając:

$$f^{-1}(C) = C^d \text{ MOD } n$$

czyli resztę z liczby C^d dzielonej przez n. Reszta z liczby C^d jest jedną z liczb pomiędzy 0 i n, przystającą do C^d modulo n. Ponieważ B^e przystaje modulo n do swojej reszty C, więc na mocy twierdzenia 3 mamy:

$$C^d \equiv (B^e)^d = B^{ed} \equiv B \pmod{n}$$

W ten sposób uzyskuje się oryginalną wiadomość. A zatem, oryginalna wiadomość B, $0 < B < n$, jest resztą z liczby C^d z dzielenia przez n.

System RSA spełnia warunki systemu jawnego klucza:

1. Liczby p i q są dużymi liczbami pierwszymi. Podobnie liczby B, e, C, d są dużymi liczbami, lecz istnieją szybkie algorytmy znajdowania reszt dla B^e i C^d modulo n. Tak więc algorytmy kodujące i dekodujące systemu RSA są rachunkowo szybkie i skuteczne.
2. Używając systemu RSA, każdy użytkownik w sieci używa komputera do wyboru odpowiednich p, q, d, a następnie do obliczenia n, k, e. Liczby e i n są publicznie udostępnione, natomiast czynniki pierwsze p i q dla n oraz liczby d i k są utajnione. Każdy użytkownik komputera może zakodować informacje używając e oraz n, ale praktycznie nie ma sposobu, dla niewtajemni-

czonych, aby określić d bez znalezienia najpierw p i q (poprzez rozłożenie n na czynniki). A zatem, system RSA jest bezpieczny dopóty, dopóki nie dysponuje się szybkimi metodami rozkładania liczb na czynniki*.

Uwaga!

Jeśli n jest wybrane jak wyżej, mogą istnieć informacje, które w formie liczbowej są większe niż n. W takich przypadkach oryginalna wiadomość jest rozkładana na kilka bloków, których wartość liczbowa jest mniejsza niż n.

3. Przykład szyfrowania za pomocą RSA

Rozważmy następującą sytuację: Odziały banku mają w sposób tajny przekazać centrali informację o wielkości dziennego zysku.

Do zaszyfrowania informacji można posłużyć się algorytmem RSA.

Niech $p = 109$, $q = 3361$. Wówczas $n = 366349$, $k = 362880$.

Niech $d = 1259$, która jest względnie pierwsza z 362880.

Rozwiązując równanie:

$$1259 \cdot x \equiv 1 \pmod{362880}$$

otrzymuje się $e = 44099$.

Zakodowana ma być wiadomość:

DZIENNY ZYSK (PLN) DRUGI MAJA BR.: 200000

przy czym umawiamy się, że wiadomość po dwukropku oznacza liczbę, której nie zamienia się na inną postać numeryczną.

Można zakodować tylko liczby mniejsze niż $n = 366349$. Tak więc wiadomość zapisuje się w blokach trzyliterowych (spację oznaczono symbolem #):

DZI	ENN	Y#Z	YSK	#(P	LN)
042609	051414	250026	251911	002916	121431

#DR	UGI	#MA	JA#	BR.	:##	200000
000418	210709	001301	100100	021827	300000	200000

Każdy blok jest liczbą mniejszą niż 366349.

Dla pierwszego bloku, 042609, oblicza się:

$$x \equiv 42609^{44099} \pmod{366349}$$

skąd $x = 141007$.

* Dysponując 100-cyfrowymi liczbami p i q, przy osiąganej przez współczesne komputery szybkości obliczeniowej, czas znalezienia rozkładu na czynniki pierwsze jest szacowany na kilkadziesiąt lat (według [2]).

Inne bloki są kodowane podobnie, a zatem zakodowana forma wiadomości jest następująca:

141007 281345 227082 122712 055921 257754
268301 191382 144634 018596 339204 152256 074226

Osoba otrzymująca wiadomość użyje $d = 1259$ do zdekodowania każdego bloku. Na przykład, aby zdekodować 141007, oblicza się:

$$a \equiv 141007^{1259} \pmod{366349}$$

skąd $a = 042609$.

Jest to oryginalny pierwszy blok D Z I.

4. Zastosowanie RSA do podpisów cyfrowych

RSA może być także stosowany do podpisów cyfrowych. Podpis cyfrowy jest dowodem wykonania transakcji, przeprowadzonej za pomocą komputera, przez konkretną osobę. Definicja podpisu cyfrowego, według polskiej normy PN-I-02000, brzmi: „Przekształcenie kryptograficzne danych umożliwiające odbiorcy danych sprawdzenie autentyczności i integralności danych oraz zapewniające nadawcy ochronę przed sfałszowaniem danych przez odbiorcę" (na podstawie: http://www.ebanki.info). W praktyce, złożenie podpisu cyfrowego polega na dołączeniu do wiadomości (transakcji) dodatkowej informacji mającej na celu weryfikację jej źródła.

Do złożenia podpisu cyfrowego jest potrzebny klucz asymetryczny (para kluczy publiczny/prywatny). Generowanie podpisu dokonuje się etapami: dane dokonywanej operacji są przekształcane w ciąg bitów, dla którego oblicza się unikatową krótką wartość, tzw. hasz*. Uzyskany hasz jest następnie szyfrowany prywatnym kluczem użytkownika za pomocą algorytmu kodującego (np. RSA), co uniemożliwia równoczesne zmiany w danych operacji i adekwatne do zmian podrobienie podpisu cyfrowego. Ten zaszyfrowany skrót stanowi podpis użytkownika, który wraz z danymi operacji jest transmitowany do serwera bankowego, gdzie następuje weryfikacja prawdziwości podpisu. Porównuje się wartości, które uzyskuje się dekodując, za pomocą klucza publicznego, otrzymany podpis oraz obliczając hasz dla otrzymanej (sformatowanej) wiadomości oryginalnej.

* Hasz uzyskuje się za pomocą tzw. funkcji haszującej, która przyporządkowuje każdej liczbie pewną wartość określaną również jako wartość skrótu funkcji. Zastosowanie algorytmu haszującego umożliwia wykrycie wszelkich modyfikacji w danych transakcji, jeżeli takowe zostałyby dokonane.

Jeżeli uzyskane wartości są identyczne, oznacza to, że podpis jest prawidłowy i operacja jest wykonywana.

Przy tworzeniu podpisów cyfrowych, w których do szyfrowania wykorzystuje się algorytm RSA, koduje się hasz wiadomości kluczem prywatnym, natomiast dekodowanie odbywa się za pomocą klucza publicznego, czyli szyfrowanie dokonuje się z odwrotnym zastosowaniem klucza.

Rozważmy przykład: Chcemy przesłać dyspozycję do banku, której sformatowana postać jest następującym ciągiem zer i jedynek (wartość umowna):

$$11000010111100000000011010111011000$$

Niech funkcja haszująca polega na obliczeniu dla danej liczby reszty z dzielenia przez 1121893, czyli $h(w) = w$ MOD 1121893. Hasz wiadomości będzie wartością:

$$11000010111100000000011010111011000 \text{ MOD } 1121893 = 821037$$

Uzyskany skrót jest szyfrowany i po dołączeniu do sformatowanej wiadomości przesyłany drogą elektroniczną do banku.

Niech kluczem prywatnym użytkownika będzie para liczb $(d, n) = (1259, 3666851)$, a kluczem publicznym para $(e, n) = (1535939, 3666851)$. Wówczas podpis dołączony do wiadomości będzie liczbą:

$$p \equiv 821037^{1259} \pmod{3666851}$$

czyli $p = 812773$.

Bankowy serwer otrzymuje wiadomość, dla której liczy hasz, czyli $h(w) = 821037$, a następnie dekoduje kluczem publicznym uzyskany podpis, licząc:

$$x \equiv 812773^{1535939} \pmod{3666851}$$

czyli $x = 821037$. Ponieważ $h(w) = x$, więc podpis jest prawidłowy i transakcja może zostać wykonana.

Podsumowanie

Narzędzia matematyczne mają szerokie zastosowanie w różnych dziedzinach nauki. Przedstawione w niniejszym artykule zastosowanie systemu RSA, powszechnie stosowanego w kryptografii, miało na celu ukazanie jednego z takich narzędzi, które dzięki powszechnie dostępnej technice komputerowej jest wykorzystywane do różnego typu transakcji internetowych. Kodowanie informacji oraz składnie podpisów cyfrowych stanowi gwarancję bezpieczeństwa dokonywanych operacji.

Algorytm (system) RSA, który jest jednym z kilku stosowanych algorytmów do różnego rodzaju operacji wykorzystujących szyfrowanie, jest pierwszym z systemów kryptografii asymetrycznej, której rozwój dał możliwość zastosowania swoich narzędzi w systemach elektronicznego uwierzytelniania, obsługi podpisów cyfrowych czy do kodowania poczty. Użyteczność tego rodzaju narzędzi w różnych dziedzinach życia jest niezaprzeczalnym argumentem za przybliżaniem stosowanych zagadnień teoretycznych.

Literatura

1. Bednarski T.: Elementy matematyki w naukach ekonomicznych. Podręcznik dla studentów ekonomii. Oficyna Ekonomiczna, Kraków 2004.
2. Bronsztejn I.N., Siemiendiajew K.A., Musiol G., Mühlig H.: Nowoczesne Kompendium Matematyki. Wydawnictwo Naukowe PWN, Warszawa 2003.
3. Hungerford T.W.: Abstract Algebra – An Introduction. Cleveland University, 1990.
4. Pośpiech E.: Utajnianie danych. Praca magisterska, Gliwice 1995.
5. Robling-Denning D.E.: Kryptografia i ochrona danych. Wydawnictwa Naukowo-Techniczne, Warszawa 1992.
6. Ross K.A., Wright C.R.B.: Matematyka dyskretna. Wydawnictwo Naukowe PWN, Warszawa 2003.

SELECTED APPLICATIONS OF THE RSA CRYPTOSYSTEM

Summary

Possibilities of using computer technology become larger and larger. Nowadays, sending e-mails, e-commerce or Internet banking are more and more popular not only among the institution representatives but also among private users. One of the most important elements connected with these activities is safety of such transactions. In this case, advanced mathematical and computer tools, such as modern cryptography, are of crucial importance.

The purpose of the article is to present one of the public-key cryptography (asymmetric cryptography) algorithms – RSA algorithm. It is based on number theory. It uses a pair of cryptographic keys – a public key (which may be distributed) and a private key (which is secret). A message encrypted with the public key can be decrypted only with the corresponding private key. The keys are related mathematically but the private key cannot be practically derived from the public key.

Public-key cryptography are used in two main branches: public-key encryption and digital signatures. The article contains examples of the two mentioned above applications of the RSA cryptosystem.

Michalina Szłapka

ALGORYTM DOBORU ZMIENNYCH OBJAŚNIAJĄCYCH O NIESTABILNEJ SILE ZALEŻNOŚCI

Wprowadzenie

W jednym ze swoich wcześniejszych artykułów autorka zauważyła, że czas wpływa na siłę zależności pomiędzy zmiennymi lub raczej siła tych zależności zmienia się w miarę upływu czasu [5]. Artykuł jest próbą wskazania pewnego sposobu postępowania w sytuacji, gdy zależność zmiennych zmieniała swoją siłę i kierunek w badanym okresie.

Problem zmienności siły zależności między zmiennymi był w literaturze poruszany kilkakrotnie. Dostępne opracowania dotyczą jednak z reguły zmian parametrów modelu ekonometrycznego, zakładając stałość zbioru zmiennych objaśniających dobranych do budowy tego modelu. W takich sytuacjach autorzy proponują m.in. wprowadzanie zmiennych zero-jedynkowych pozwalających na zmianę wartości parametrów w sytuacji działania jednego lub więcej czynników niemierzalnych powodujących skokowe zmiany niektórych lub wszystkich parametrów modelu czasie [1].

Nieco inne podejście przyjęli Grabinski, Zeliaś i Wydymus w pracy [2]. Rozpatrując dobór zmiennych dla celów prognostycznych, wspominają oni, że wartości wskaźników pojemności integralnej* mogą podlegać określonym trendom w badanym czasie. W związku z tym autorzy proponują podział próby na ruchome okresy o ustalonej stałej liczebności s (przy czym dopuszczają również stosunkowo niewielką wartość s = 5). Pierwszy z utworzonych okresów obejmowałby obserwacje 1-s, następny 2-(s+1) itd. aż do wyczerpania całego zbioru

* Metoda Hellwiga. Punkt 3.1 artykułu.

obserwacji. Dla w ten sposób ustalonych okresów autorzy ci proponują wyznaczanie macierzy współczynników korelacji zmiennych oraz wartości wskaźników Hellwiga dla wszystkich możliwych kombinacji. Zaletą tego podejścia jest uwzględnienie zmian zależności zmiennych w czasie, jednak ma ono kilka wad. Pierwszą z nich jest trudność interpretacji uzyskanych wyników. Przyjęty sposób wyznaczania okresów dla obliczeń współczynników korelacji i wskaźników pojemności informacyjnej klasyfikuje określone obserwacje jednocześnie do kilku okresów, przez co np. dla drugiej obserwacji zmiennych trudno powiedzieć, jak kształtowała się w niej zależność między zmiennymi lub która ich kombinacja była najlepsza, ponieważ należała ona zarówno do okresu 1-s, jak i 2- (s+1) i być może kilku innych. W efekcie zaproponowane podejście może służyć tylko i wyłącznie do uzyskiwania informacji o kształtowaniu się zależności w najbliższych s okresach. Dodatkową wadą jest duża automatyczność opisanego sposobu postępowania – zmienne są dzielone na podokresy bez przeprowadzanej analizy charakteru ich zależności w kolejnych okresach. Kolejnym utrudnieniem zaproponowanej w tej pracy metody jest jej duża pracochłonność. Zalecane przez autorów jest wyznaczenie kilkunastu macierzy współczynników korelacji i prognozowanie wartości wskaźników dla każdej kombinacji, podczas gdy wystarczające wydaje się zbadanie trendu samych miar zależności, określenie ich przyszłych wartości i dopiero dla nich wyznaczenie przyszłej najlepszej kombinacji zmiennych.

W niniejszym artykule przeprowadzono badanie zmian zależności zmiennych poprzez określenie, jak długo między wybranymi dwoma zmiennymi utrzymywała się ta sama siła zależności. W tym celu wykorzystano wzrokową ocenę charakteru zależności na podstawie diagramu korelacyjnego. Dla okresów podobnego kształtowania się zależności obliczono współczynniki korelacji. Następnie dla każdej z analizowanych jednostek czasu wyznaczono macierz współczynników zależności. Na podstawie tej macierzy wyodrębniono najlepszą kombinację zmiennych objaśniających (dla każdej jednostki czasowej, w której zależność miała zbliżony charakter założono tę samą wartość miernika siły zależności), przy użyciu metody Hellwiga.

Analiza przeprowadzana w niniejszym artykule ma na celu zbadanie, jakie są możliwe zmiany zależności zmiennych w miarę upływu czasu oraz jak zmieniają się kombinacje zmiennych najistotniej kształtujących zmienną objaśnianą.

1. Prezentacja materiału liczbowego

W pracy zastosowano materiał liczbowy dotyczący liczby dzieci przebywających w domach dziecka oraz czterech zmiennych przypuszczalnie wpływających na tę wielkość. Zebrane dane pochodzą z Rocznika Statystycznego i dotyczą sytuacji w Polsce w latach 1986-2000.

Zebrana próba jest stosunkowo niewielka, podobnie jak utworzony zbiór potencjalnych zmiennych objaśniających. Wynika to z faktu, że przedstawiony materiał stanowi jedynie pretekst do zaprezentowania pewnego podejścia do doboru zmiennych i badania zmian ich wzajemnej zależności. Celem autorki było zachowanie możliwie największej czytelności i prostoty ilustrujących przyjęte postępowanie rysunków oraz obliczeń.

Tabela 1

Liczba wychowanków domów dziecka i kształtujące ją czynniki w latach 1986-2000

Lata	Lp.	y	X_1	X_2	X_3	X_4
1986	1	16975	37572	50580	376,3	2,217
1987	2	16659	37764	49707	378,4	2,154
1988	3	16130	37885	48650	370,8	2,126
1989	4	16600	38038	47200	381,1	2,078
1990	5	15370	38183	42436	390,3	2,039
1991	6	15018	38245	33823	404	2,049
1992	7	15113	38365	32024	393,1	1,992
1993	8	15397	38459	27891	390,9	1,847
1994	9	17491	38544	31574	386,4	1,798
1995	10	18678	38588	38115	386,1	1,611
1996	11	18766	38618	39449	385,5	1,58
1997	12	18712	38650	42549	380,2	1,513
1998	13	18508	38666	45230	375,3	1,431
1999	14	18101	38654	42020	381,4	1,366
2000	15	18225	38646	42770	368	1,337

Źródło: [4].

Jako zmienną objaśniającą y przyjęto liczbę wychowanków domów dziecka. Potencjalne zmienne objaśniające zostały oznaczone następująco:

- X_1 – liczba ludności Polski ogółem,
- X_2 – liczba małżeństw rozwiązanych przez rozwody,
- X_3 – liczba zgonów ludności,
- X_4 – współczynnik dzietności (liczba dzieci urodzonych przeciętnie przez kobietę w wieku 15-49 lat).

Dla tak zdefiniowanych zmiennych w dalszej części artykułu przeprowadzono analizę siły i charakteru zależności zmiennych na podstawie wzrokowej oceny diagramów korelacyjnych.

2. Wnioskowanie o zmianie zależności na podstawie diagramów korelacyjnych

Diagramy korelacyjne stanowią najprostszy sposób zobrazowania zależności dwóch zmiennych. Ich istotą jest zaznaczenie w dwuwymiarowej przestrzeni punktów o współrzędnych odpowiadających kolejnym wartościom badanych cech. Stosowane są w celu określenia charakteru zależności zmiennych. Ponadto analiza rozrzucenia punktów w stosunku do ich linii trendu pozwala na dokonanie wstępnej wzrokowej analizy siły zależności zmiennych.

Powszechnie przyjmuje się, że pomiędzy zmiennymi można przyjąć zależność liniową, gdy punkty diagramu otoczy się elipsą. Im „węższa" jest narysowana elipsa, tym silniejsza powinna być liniowa zależność zmiennych. Przyjmuje się również, że rozkład punktów mający kształt kolisty lub też przypominający podkowę o końcach zwróconych w stronę rosnących wartości osi OX świadczy o braku zależności badanych zmiennych.

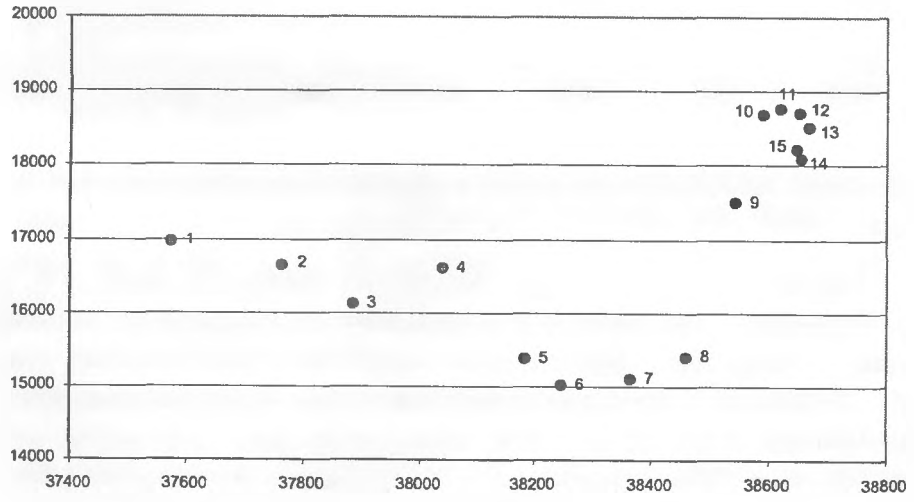
Zaproponowane w dalszej części artykułu wnioskowanie na podstawie diagramów korelacyjnych uwzględnia wymienione powyżej zasady, lecz wykorzystuje dodatkowo możliwość zmiany charakteru zależności zmiennych w miarę upływu czasu. Poszerzając diagramy korelacyjne o numer obserwacji, którą ilustruje dany punkt diagramu, można wyróżnić w ramach badanej próby różne zależności liniowe charakteryzujące kilka chronologicznie po sobie następujących obserwacji. Niejednokrotnie wnioskowanie bez określenia okresów, z których pochodziły dane, prowadziło do stwierdzenia słabej zależności zmiennych lub jej braku.

Przytoczone wnioskowanie można dokładniej przeanalizować na podstawie poniższych rysunków dotyczących zaprezentowanego w punkcie 1 materiału liczbowego.

Dla odróżnienia okresów, w których zależność zmiennych miała różny charakter, zaznaczono je różnymi kolorami.

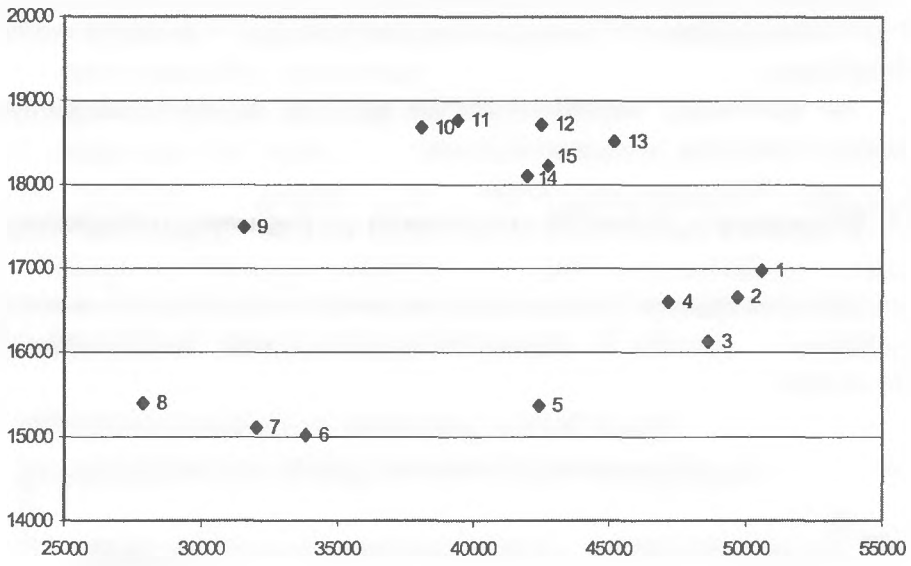
2.1. Diagramy zależności zmiennych ze zmienną objaśnianą

Poniższe diagramy obrazują rozkłady zmiennych objaśnianych przyjętych w punkcie 1 w stosunku do zmiennej objaśnianej – liczby wychowanków domów dziecka.



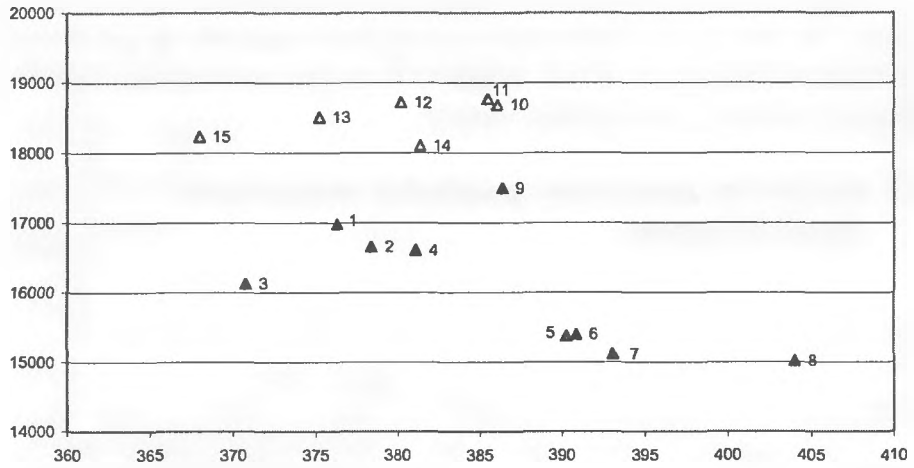
Rys. 1. Diagram korelacyjny zmiennej y (liczba wychowanków domów dziecka) oraz zmiennej X_1 (liczba ludności ogółem)

Analizując powyższy diagram, można podzielić obserwacje zmiennych na dwa okresy. Obserwacje od 1-8 wykazują dość silną malejącą zależność zmiennych. W związku z rozkładem punktów w tych okresach przeprowadzono badanie siły związku krzywoliniowego zmiennych, które wskazało na dopuszczalne przyjęcie zależności liniowej. Okresy od 10-15 wskazują na znaczne osłabienie siły zależności zmiennych. Problemowe okazało się przydzielenie obserwacji 9, do którejś z wymienionych grup. Autorka zdecydowała się potraktować tę obserwację osobno, zakładając dla niej zerową siłę zależności ze zmienną y.



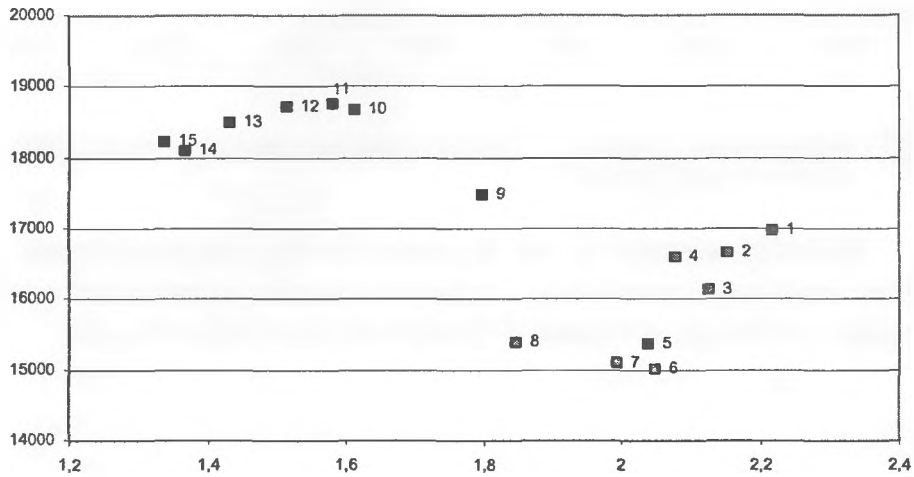
Rys. 2. Diagram korelacyjny zmiennej y (liczba wychowanków domów dziecka) oraz zmiennej X_2 (liczba małżeństw rozwiązanych przez rozwody)

Dla powyższych zmiennych wyodrębniono analogicznie okresy 1-8, 10--15. Obserwację 9 potraktowano jako odstępstwo od występujących zależności zmiennych. Dodatkowo przeprowadzono analizę siły krzywoliniowości zmiennych w okresach 1-8. Otrzymane wyniki wskazywały na przyjęcie tego charakteru zależności, w związku z czym w wymienionych okresach do analizy zależności włączono również zmienną $Z_1 = X_2^2$. Zmienne wskazane przez przyjętą metodę doboru zmiennych objaśniających jako najlepsze były jednak takie same, jak przy przyjęciu zależności liniowej zmiennych. Zmienna Z_1 stanowiła przekształcenie X_2, zatem jej pojawienie się w najlepszej kombinacji potraktowano jako potwierdzenie siły wpływu zmiennej X_2 na y. Stąd w artykule pominięto obliczenia dla tej zmiennej pomocniczej. Obserwacja 9 okazała się być trudna do sklasyfikowania, w związku z czym została potraktowana jako odstępstwo od przyjętych zależności. Dla tego momentu próby siłę związku y-X_2 uznano za zerową.



Rys. 3. Diagram korelacyjny zmiennej y (liczba wychowanków domów dziecka) oraz zmiennej X_3 (liczba zgonów ogółem)

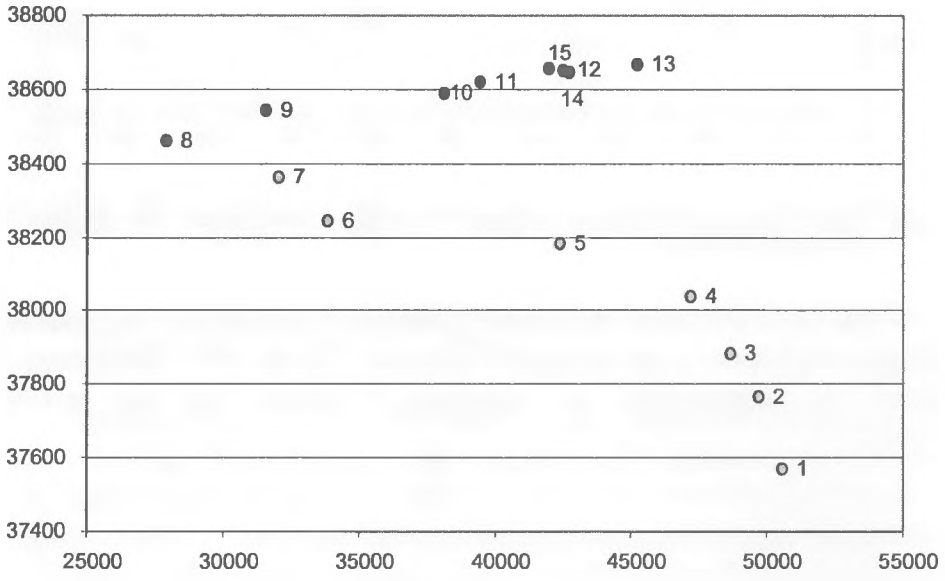
Tak jak w przypadku poprzednich zmiennych, stwierdzono, że zależność zmiennych kształtowała się podobnie w okresach 1-8 oraz 10-15. Obserwację 9 uznano za nieprzystającą do pozostałych i przyjęto dla niej wartość $r_3 = 0$.



Rys. 4. Diagram korelacyjny zmiennej y (liczba wychowanków domów dziecka) oraz zmiennej X_4 (współczynnik dzietności)

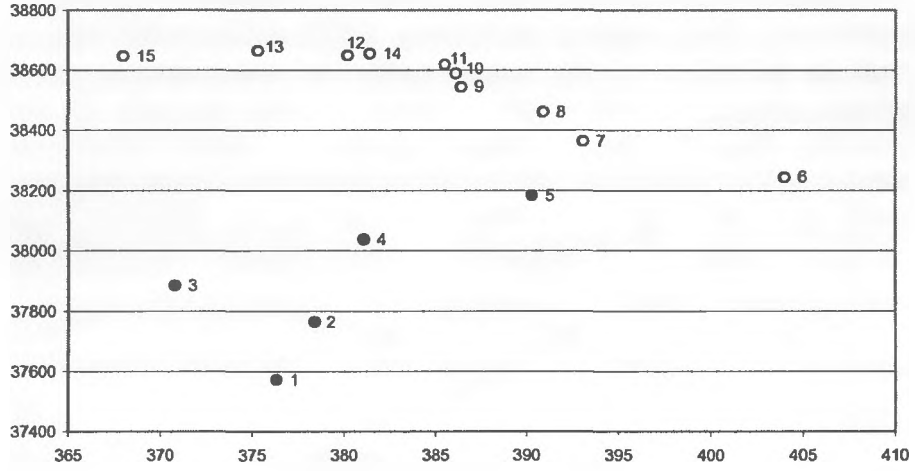
Dla powyższych zmiennych badaną próbę podzielono również na dwa okresy: 1-8 oraz 10-15. Obserwacja 9 wyraźnie nie pasowała do kształtowania się zależności zmiennych w tych okresach. W obydwu analizowanych okresach zależność zmiennych ma charakter rosnący.

2.2. Diagramy zależności pomiędzy zmiennymi objaśniającymi



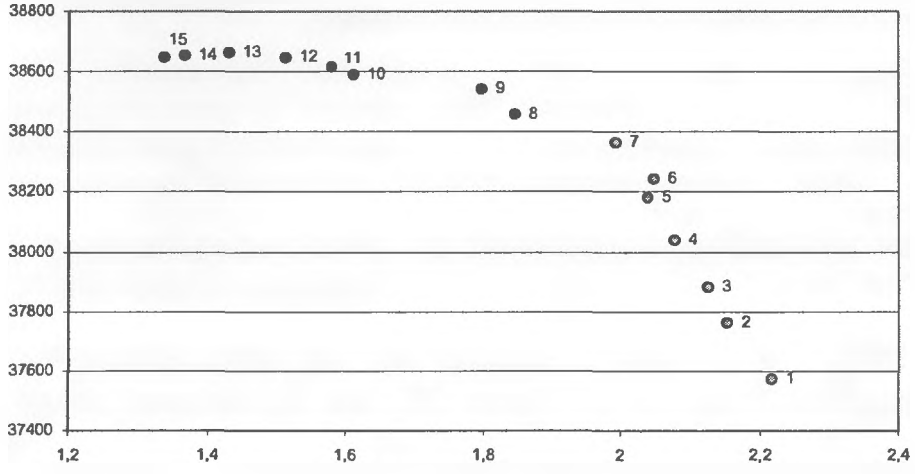
Rys. 5. Diagram korelacyjny zmiennej X_1 (ludność ogółem) oraz zmiennej X_2 (liczba małżeństw rozwiązanych przez rozwody)

Zależność zmiennych X_1 oraz X_2 wskazuje na dwa przeciwne kierunki zależności zmiennych. Obserwacje 1-7 charakteryzują się silną liniową zależnością malejącą, podczas gdy w okresach 8-15 zależność zmiennych jest rosnąca.



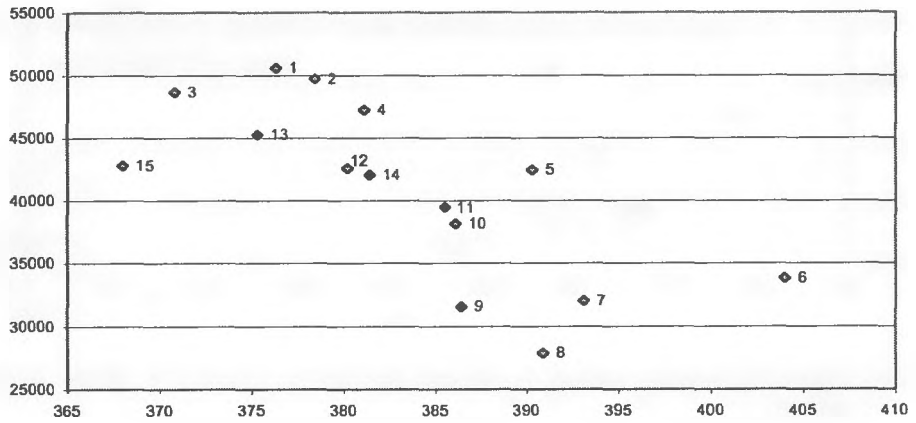
Rys. 6. Diagram korelacyjny zmiennej X_1 (ludność ogółem) oraz zmiennej X_3 (liczba zgonów ogółem)

W przypadku zależności zmiennych X_1 oraz X_3 można wyróżnić dwa okresy badanej próby. Zmienne pochodzące z obserwacji 1-5 charakteryzują się słabą zależnością rosnącą. Punkty diagramu dotyczące okresu 6-15 świadczą natomiast o silnej liniowej malejącej zależności zmiennych.



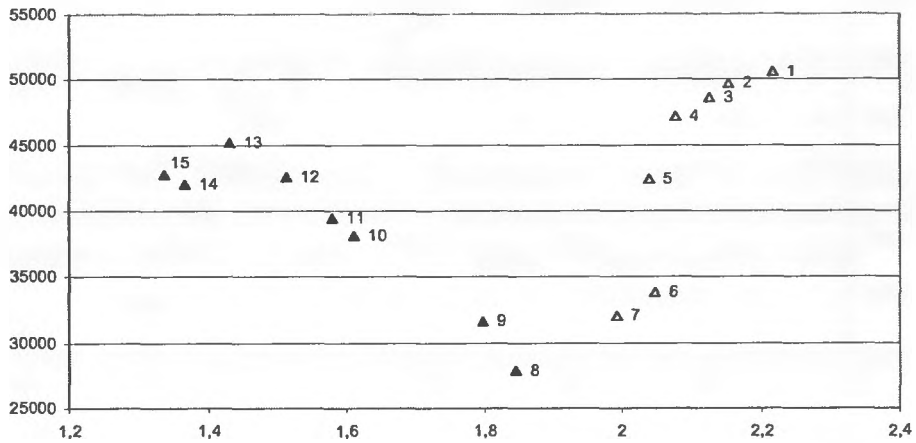
Rys. 7. Diagram korelacyjny zmiennej X_1 (ludność ogółem) oraz zmiennej X_4 (współczynnik dzietności)

Przeprowadzony dla powyższych zmiennych podział próby na okresy: 1-7 oraz 8-15 miał na celu jedynie sprowadzenie charakteru zależności zmiennych do bardziej liniowego. Dla całej badanej próby zależność zmiennych ma silny charakter malejący.



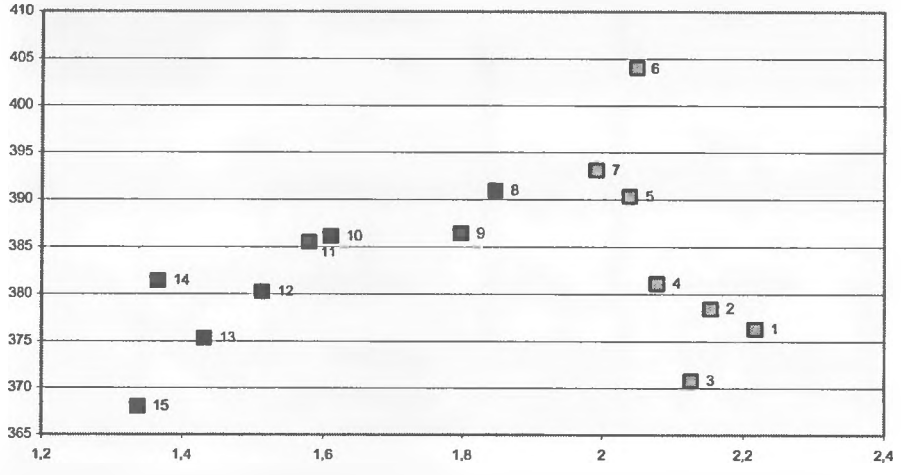
Rys. 8. Diagram korelacyjny zmiennej X_2 (liczba małżeństw rozwiązanych przez rozwody) oraz zmiennej X_3 (liczba zgonów ogółem)

Dla zmiennych X_2 oraz X_3 wyróżniono dwa okresy podobnego przebiegu ich zależności: 1-6, 7-15. W obydwu z tych okresów zależność zmiennych miała charakter malejący. Dla wcześniejszych zależności zależność ta jest nieco silniejsza i wyraźniej widoczna niż dla ostatnich ośmiu.



Rys. 9. Diagram korelacyjny zmiennej X_2 (liczba małżeństw rozwiązanych przez rozwody) i zmiennej X_4 (współczynnik dzietności)

W przypadku diagramu korelacyjnego tych dwóch zmiennych problemem było określenie, do której z wyróżnionych grup powinna należeć obserwacja 8, ponieważ jest ona zgodna z obydwoma wyróżnionymi kierunkami zależności. Gdyby nie dokonana analiza okresów, z których pochodzą punkty diagramu badanej próby, można by wyciągnąć wniosek o braku zależności zmiennych, podczas gdy numerowanie punktów umożliwiło wyróżnienie dwóch okresów o dość wyraźnych, chociaż przeciwnie ukierunkowanych liniowych zależnościach zmiennych. Wyróżniono dwa okresy jednolitej zależności: 1-7 oraz 8-5.



Rys. 10. Diagram korelacyjny zmiennej X_3 (liczba zgonów ogółem) oraz zmiennej X_4 (współczynnik dzietności)

Na podstawie powyższego diagramu wyznaczono dla zmiennych X_3 oraz X_4 następujące okresy ich jednolitego kształtowania się: 1-7 oraz 8-15. Pierwszy z wyszczególnionych okresów charakteryzuje się malejącą zależnością zmiennych, podczas gdy od 8 obserwacji zależność ta przybrała charakter rosnący.

2.3. Podział badanej próby na podokresy o jednakowej sile i kierunku zależności

W związku z podziałem próby badawczej w punkcie 2.2 można zauważyć, że dla wielu zmiennych okresy, na które podzielono próbę, mają różną długość.

W związku z faktem, że cała macierz współczynników będąca podstawą przeprowadzanego doboru zmiennych dotyczy jednej, tej samej liczby obserwacji, początkową próbę zmiennych należy podzielić uwzględniając okresy wyróżnione dla każdej z badanych par zmiennych.

Zastosowany sposób postępowania przedstawia tabela 2.

Tabela 2

Podział 15-elementowej próby na podstawie okresów jednolitego kształtowania się zależności poszczególnych par zmiennych

Numer obserwacji	Współczynniki korelacji			
1	r_1, r_2, r_3, r_4	$r_{12}, r_{14}, r_{24}, r_{34}$	r_{23}	r_{13}
2				
3				
4				
5				
6				
7				
8				
9	r_1, r_2, r_3, r_4			
10	r_1, r_2, r_3, r_4	$r_{12}, r_{14}, r_{24}, r_{34}$	r_{23}	r_{13}
11				
12				
13				
14				
15				

Najkrótszy z wyróżnionych w punkcie 2.2 w ramach badanej próby okresów dotyczył obserwacji 1-5, więc dla tego okresu zbudowano pierwszą z macierzy miar zależności między zmiennymi. Inne wyróżnione okresy początkowe to: 1-6, 1-7, 1-8. Stąd macierz współczynników zależności zmiennych będzie inna dla obserwacji 6, 7 oraz 8 badanej próby. Dla obserwacji 10-15 każdy wyznaczony współczynnik nie zmieni już swojej wartości, dlatego dobór zmiennych dla tego okresu zostanie przeprowadzony na podstawie jednej macierzy.

Tak wyróżnione macierze będą zawierały współczynniki korelacji wyznaczone dla wyróżnionych w punkcie 2.2 i przedstawionych w tabeli 1 okresów, w których zawiera się dany numer obserwacji.

Ze względu na niemożność zaklasyfikowania w diagramach 1-4 obserwacji 9 do okresów o jednolitej sile i kierunku zależności zmiennych, przyjęto dla niej zerową wartość miernika zależności. W związku z faktem, że sytuacja ta dotyczyła związku wszystkich zmiennych objaśniających ze zmienną objaśnianą, dla obserwacji 9 niemożliwe okazało się przeprowadzenie procedury doboru zmiennych objaśniających.

3. Procedura doboru zmiennych objaśniających

3.1. Zastosowane mierniki zależności i przyjęta metoda doboru zmiennych

Podstawą doboru zmiennych są macierze współczynników zależności między zmiennymi:

$$R_0 = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix}, \qquad R = \begin{bmatrix} 1 & r_{12} & \cdots & r_{1n} \\ r_{21} & 1 & \cdots & r_{2n} \\ \vdots & \vdots & & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nn} \end{bmatrix} \tag{1}$$

gdzie:

R_0 – wektor współczynników korelacji zmiennych objaśniających ze zmienną objaśnianą,

R – macierz współczynników korelacji między zmiennymi objaśniającymi,

r_i – współczynnik korelacji pomiędzy zmienną objaśnianą a i-tą potencjalną zmienną objaśniającą,

r_{ij} – współczynnik korelacji pomiędzy zmienną objaśniającą Xi a zmienną objaśniającą Xj.

Jako miernik siły zależności między zmiennymi przyjęto współczynnik korelacji liniowej Pearsona najczęściej stosowany do pomiaru siły związku między dwoma zmiennymi mierzonymi na skali nominalnej:

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \cdot \sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{2}$$

gdzie:

n – liczebność zebranej próby obserwacji,

i = 1, 2, 3,.....n – numer obserwacji zmiennej,

X, Y – zmienne, pomiędzy którymi jest badana siła zależności.

W artykule posłużono się metodą wskaźników pojemności informacyjnej Hellwiga jako jedną z najczęściej stosowanych metod doboru zmiennych, cechującą się dużą prostotą wykonywanych obliczeń.

W metodzie tej kryterium doboru zmiennych jest maksymalizacja integralnego wskaźnika pojemności informacyjnej (H_l) wyznaczanego dla każdej możliwej kombinacji zmiennych objaśniających.

Liczba rozpatrywanych kombinacji wynosi:

$$L = 2^m - 1 \tag{3}$$

gdzie m oznacza liczbę potencjalnych zmiennych objaśniających.

Integralny wskaźnik pojemności informacyjnej wyraża się wzorem:

$$H_l = \sum_{j=1}^{k} h_{lj} \tag{4}$$

$$h_{lj} = \frac{r_j^2}{1 + \sum_{\substack{i=1 \\ i \neq j}}^{k} |r_{ij}|} \tag{5}$$

gdzie:

l = 1,2,...L – numer rozpatrywanej kombinacji zmiennych,

j = 1,2,....k – numer zmiennej w kombinacji,

k – liczba zmiennych w rozpatrywanej kombinacji,

h_{lj} – indywidualny wskaźnik pojemności informacyjnej zmiennych w ramach rozpatrywanej kombinacji.

Ponieważ pomiędzy wyodrębnionymi powyższą metodą zmiennymi może zachodzić kataliza zniekształcająca otrzymane wyniki, stąd dodatkowo w pracy zostanie wykorzystany test na występowanie tego efektu.

$$|r_{ij}| > min\left\{\left|\frac{r_i}{r_j}\right|; \left|\frac{r_j}{r_i}\right|\right\} \tag{6}$$

gdzie r_{ij} to wzajemne skorelowanie i-tej i j-tej zmiennej objaśniającej.

W przypadku gdy zapisany warunek jest spełniony ze zmiennych X_i oraz X_j, do kombinacji można przyjąć tylko tę, która była silniej skorelowana ze zmienną objaśnianą Y.

3.2. Wyniki zastosowanej procedury

Dla określonych w punkcie 2 okresów badania zależności między zmiennymi, zgodnie z zasadami opisanymi w punkcie 2.3, otrzymano następujące współczynniki zależności między zmiennymi dla kolejnych obserwacji:

– okresy 1-5: $R=\begin{bmatrix} 1 & -0{,}896 & 0{,}6756 & -0{,}9918 \\ -0{,}896 & 1 & -0{,}9558 & 0{,}8489 \\ 0{,}6756 & -0{,}9558 & 1 & -0{,}7235 \\ -0{,}9918 & 0{,}8489 & -0{,}7235 & 1 \end{bmatrix}; R_0=\begin{bmatrix} -0{,}886 \\ 0{,}8633 \\ -0{,}8383 \\ 0{,}7357 \end{bmatrix}$

– okres 6: $R=\begin{bmatrix} 1 & -0{,}896 & -0{,}8898 & -0{,}9918 \\ -0{,}896 & 1 & -0{,}9558 & 0{,}8489 \\ -0{,}8898 & -0{,}9558 & 1 & -0{,}7235 \\ -0{,}9918 & 0{,}8489 & -0{,}7235 & 1 \end{bmatrix}; R_0=\begin{bmatrix} -0{,}886 \\ 0{,}8633 \\ -0{,}8383 \\ 0{,}7357 \end{bmatrix}$

– okres 7: $R=\begin{bmatrix} 1 & -0{,}896 & -0{,}8898 & -0{,}9918 \\ -0{,}896 & 1 & -0{,}8062 & 0{,}8489 \\ -0{,}8898 & -0{,}8062 & 1 & -0{,}7235 \\ -0{,}9918 & 0{,}8489 & -0{,}7235 & 1 \end{bmatrix}; R_0=\begin{bmatrix} -0{,}886 \\ 0{,}8633 \\ -0{,}8383 \\ 0{,}7357 \end{bmatrix}$

– okres 8: $R=\begin{bmatrix} 1 & 0{,}9806 & -0{,}8898 & -0{,}9151 \\ 0{,}9806 & 1 & -0{,}8062 & -0{,}9312 \\ -0{,}8898 & -0{,}8062 & 1 & 0{,}8402 \\ -0{,}9151 & -0{,}9312 & 0{,}8402 & 1 \end{bmatrix}; R_0=\begin{bmatrix} -0{,}886 \\ 0{,}8633 \\ -0{,}8383 \\ 0{,}7357 \end{bmatrix}$

– okres 9: $R=\begin{bmatrix} 1 & 0{,}9806 & -0{,}8898 & -0{,}9151 \\ 0{,}9806 & 1 & -0{,}8062 & -0{,}9312 \\ -0{,}8898 & -0{,}8062 & 1 & 0{,}8402 \\ -0{,}9151 & -0{,}9312 & 0{,}8402 & 1 \end{bmatrix}; R_0=\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

– okresy 10-15: $R=\begin{bmatrix} 1 & 0{,}9806 & -0{,}8898 & -0{,}9151 \\ 0{,}9806 & 1 & -0{,}8062 & -0{,}9312 \\ -0{,}8898 & -0{,}8062 & 1 & 0{,}8402 \\ -0{,}9151 & -0{,}9312 & 0{,}8402 & 1 \end{bmatrix}; R_0=\begin{bmatrix} -0{,}5038 \\ -0{,}4126 \\ 0{,}5598 \\ 0{,}9045 \end{bmatrix}$

Po podstawieniu do wzorów (1)(-4), dla wymienionych obserwacji, maksymalną wartość wskaźnika pojemności informacyjnej otrzymano dla kombinacji przedstawionych w tabeli 2.

Tabela 3

Wyniki metody doboru zmiennych Hellwiga dla wydzielonych okresów próby badawczej

Numery obserwacji	Zmienne wchodzące w skład najlepszej kombinacji	Wartość H_l
1-5	$\{X_1X_3\}$	0,8879
6	$\{X_1X_2\}$	0,8228
7	$\{X_1X_2X_3\}$	0,8282
8	$\{X_2X_3\}$	0,8196
9	brak	brak
10-15	$\{X_4\}$	0,8182

Jak widać w tabeli 3, rezultaty metod doboru zmiennych zależały w dużym stopniu od okresów badania współczynników korelacji zmiennej objaśnianej ze zmiennymi objaśniającymi, ponieważ podział na te właśnie okresy znajduje odzwierciedlenie w uzyskanych wynikach. Zmienne o numerach obserwacji 1-8 to zmienne zawierające w prawie wszystkich swoich kombinacjach zmienną X_1 – liczbę ludności ogółem i X_2 – liczbę rozwodów lub X_3 – liczbę zgonów. Dla obserwacji 10-15 najlepszą kombinacją zmiennych objaśniających była zmienna X_4 – współczynnik dzietności. Uzyskane w pierwszych okresach wyniki są jednak w niewielkim stopniu zróżnicowane przez zmiany zależności między zmiennymi objaśniającymi.

Dla każdej z wyróżnionych kombinacji przeprowadzono badanie efektu katalizy za pomocą wzoru 6. Wynik był negatywny, co oznacza, że na uzyskaną wysoką wartość wskaźnika nie miało wpływu skorelowanie zmiennych ze sobą.

Podsumowanie i wnioski

Przeprowadzona w artykule analiza czasowych zmian zależności między zmiennymi potwierdziła duży wpływ próby badawczej na wielkość charakteryzujących tę zależność mierników. Na zamieszczonych diagramach można zaobserwować, że często włączenie lub niewłączanie którejś z obserwacji do badanego okresu kształtowania się zależności zmiennych mogło doprowadzić do zmiany siły, a nawet kierunku wyznaczanej dla tego okresu korelacji.

Przyjęty w przedstawionym algorytmie postępowania sposób wykonywania obliczeń może stanowić pewne utrudnienie, jeśli chodzi o modelowanie zmiennych, chociażby ze względu na zmniejszenie liczebności badanej próby

przez podział jej na mniejsze okresy. Natomiast zaletą przyjętego podejścia jest wyznaczanie siły zależności dla zmiennych w okresach, o których wiadomo, że charakteryzowały się jednolitą siłą i kierunkiem zależności, co pozwala na uzyskanie dla konkretnej wybranej obserwacji jednoznacznej informacji o sile zależności zmiennych objaśniających z objaśnianą i między sobą oraz o najlepszej ich kombinacji.

W związku z przeprowadzonymi badaniami pojawia się kilka nowych pytań dotyczących metod doboru zmiennych przy niestabilnej sile łączących je zależności. Jedno z nich brzmi: „do jakiego stopnia można mówić o zmiennym charakterze zależności, a kiedy powinno się stwierdzić, że zależność między zmiennymi jest po prostu słaba?".

Wspomniany problem wymaga określenia minimalnej liczebności próby, dla której interpretacja wyznaczonych miar zależności ma sens. W niniejszym artykule autorka określiła liczebność okresów na podstawie wzrokowej oceny zbieżności kolejnych obserwacji na diagramach korelacyjnych. Uzyskana w ten sposób próba obejmowała 5-8 obserwacji, była więc dość nieliczna. Nie była jednak mniejsza niż zaproponowana w jednej z przytoczonych pozycji literaturowych.

Kolejnym problemem wynikającym z przeprowadzonej analizy jest klasyfikacja punktów, które mniej lub bardziej wyraźnie nie mieszczą się w analizowanych skupiskach punktów, jak np. dziewiąta obserwacja diagramu zależności zmiennej y ze zmiennymi objaśniającymi. Autorka zdecydowała się potraktować wymienioną obserwację jako nieprzystającą do występujących w badanych okresach zależności zmiennych. Również w przypadku badania zależności pomiędzy kilkoma innymi zmiennymi określenie, które okresy można uznać za okresy występowania pewnej jednolitej zależności zmiennych, okazało się trudne. Stąd pewną wadą zaproponowanej metodologii jest fakt, że przeprowadzona klasyfikacja jest w znacznej mierze obarczona subiektywizmem.

Rozwiązaniem obydwu wspomnianych trudności może się okazać poszerzenie klasyfikacji próby na okresy jednolitej zależności o śledzenie współczynników kierunkowych prostych przechodzących kolejno przez coraz większą liczbę punktów lub analiza zmienności punktów względem łączącej je prostej. Również wiedza merytoryczna dotycząca kształtowania się badanych zjawisk w analizowanym okresie może pomóc rozstrzygnąć kwestię, czy kształt diagramu korelacyjnego świadczy o zmianie kierunku zależności, czy raczej oznacza, iż między zmiennymi występowała bardzo słaba zależność lub nie było jej w ogóle.

Literatura

1. Ekonometryczne modele rynku. Analiza. Prognozy. Symulacja. Tom 1: Metody ekonometryczne. Red. W. Welfe. PWE, Warszawa 1977.
2. Grabiński T., Wydymus S., Zeliaś A.: Metody doboru zmiennych w modelach ekonometrycznych. PWN, Warszawa 1982.
3. Metody statystycznej analizy wielowymiarowej w badaniach marketingowych. Red. E. Gatnar, M. Walesiak. Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław 2004.
4. Rocznik demograficzny. GUS, Warszawa 2003.
5. Szłapka M.: Czas jako czynnik wpływający na dobór zmiennych objaśniających. Zeszyty Naukowe Politechniki Śląskiej, Seria: Organizacja i Zarządzanie, z. 20/1, Gliwice 2004.

ALGORITHM OF CHOOSING EXPLANATORY VARIABLES OF THE UNSTABLE RELATIONSHIP POWER

Summary

The article is a continuation of the research begun in author's earlier work. It deals with the problem of changes that affect relationships between explanatory variables themselves and between explanatory variables and the explained variable. It means that in case of using time variables, different values of correlation coefficients may be obtained for different lengths of sample and for different periods taken from the sample.

Considering this fact there is an algorithm of choosing variables method presented in this article on exemplary data. Four explanatory variables are used to explain number of children in refuges. These are: the general number of people, a number of divorces, the general number of demises, and a number of children per woman. This sample includes fifteen observations from years 1986-2000.

In order to observe when a change of the correlation direction or strength took place, ten correlation diagrams were built. This kind of visual presentation contains points which coordinate corresponding adequate values of the analysed variables. In this article there were also labels of time marked on the diagrams which helped to distinguish for which observation the relationship could be described with one measure of correlation. After the correlation coefficients had been calculated, the sample was divided in order to provide one value of correlation coefficient in every part of it. In this article the sample was divided into five parts containing observations number 1-5, 6, 7, 8 and 10-15.The ninth observation was an exception from the relationships between explanatory and explained variables, so it was excluded from further research. Then for each part of the sample the best combination of variables was indicated by applying Hellwig method.

Final results proved the necessity of considering relationships' changes in the methods of choosing variables. This research revealed some difficulties and new problems that have to be investigated. They deal with the classification of the correlation diagrams' points and minimum number of the samples which justify interpreting the measures of the variables' relationships.

Joanna Trzęsiok

O REGULARYZACJI WYBRANYCH NIEPARAMETRYCZNYCH MODELI REGRESJI

Wprowadzenie

W dobie dynamicznie rozwijających się narzędzi informatycznych mamy do czynienia z wieloma nowymi propozycjami metod regresji. Obecnie techniki numeryczne pozwalają nam na tworzenie nieparametrycznych modeli regresji, w których nieznana jest a priori analityczna postać funkcji regresji. W miarę potrzeb do modelu dołącza się funkcje składowe, których parametry są szacowane w kolejnych krokach algorytmu. Możliwość tworzenia modelu regresyjnego w sposób iteracyjny, poprzez dokładanie kolejnych funkcji bazowych oraz poprawianie jakości tego modelu, prowadzi jednak często do nadmiernego dopasowania modelu do danych ze zbioru uczącego (ang. overfitting). Uzyskujemy przez to duże błędy prognoz na zbiorze rozpoznawanym, złożonym z nowych obiektów nieuczestniczących w procesie budowy modelu.

Zachodzi konieczność kontrolowania, by wartości błędu resubstytucji, odzwierciedlające dopasowanie modelu na zbiorze uczącym, nie były zbyt małe. Nawet najmniejsze wartości tego błędu nie gwarantują równie małych wartości błędów predykcji. Przeciwnie, małe błędy na zbiorze uczącym zazwyczaj uzyskuje się dla modeli charakteryzujących się dużą złożonością i tym samym dających duże błędy prognoz na zbiorze testowym. Zachodzi więc potrzeba uzyskania pewnego kompromisu pomiędzy dopasowaniem funkcji regresji f do obserwacji ze zbioru uczącego a stopniem złożoności modelu. Kompromis ten będziemy nazywać **regularyzacją** (ang. regularization) i możemy go zapisać w postaci zadania minimalizacji funkcjonału $PRSS(f, \lambda)$ (ang. penalized residual sum of squares):

$$PRSS(f,\lambda) = RSS(f) + \lambda J(f) \rightarrow \min \tag{1}$$

gdzie $PRSS(f)$ jest miarą dopasowania modelu do danych ze zbioru uczącego (najczęściej w przypadku regresji jest to suma kwadratów reszt), natomiast $J(f)$ odzwierciedla złożoność modelu (im bardziej złożony model, tym większe uzyskujemy wartości $J(f)$). Parametr $\lambda \geq 0$, nazywany współczynnikiem kary (ang. penalty), który określa proporcje pomiędzy składowymi funkcjonału $PRSS(f,\lambda)$.

Problem regularyzacji został w różny sposób rozwiązany w nieklasycznych metodach regresji. Celem tego artykułu jest przedstawienie przykładowych rozwiązań tego zagadnienia w wybranych nieparametrycznych metodach regresji, takich jak: metoda krzywych sklejanych, MARS i MART.

1. Metoda krzywych sklejanych

Metoda krzywych sklejanych została szczegółowo omówiona w pracy [9]. Tutaj zajmiemy się głównie problemem regularyzacji modelu opartego na krzywych sklejanych.

Metoda krzywych sklejanych polega na podziale dziedziny zmiennej objaśniającej X na K rozłącznych przedziałów za pomocą uporządkowanego zbioru punktów:

$$\{\xi_i\}_{i=1,\ldots,K-1}$$

nazywanych węzłami. W każdym z powstałych przedziałów $\langle \xi_i, \xi_{i+1} \rangle$ szukamy, posługując się metodą najmniejszych kwadratów, funkcji regresji fi, która jest wielomianem co najwyżej stopnia S.

W najprostszym przypadku funkcje fi mogą być funkcjami stałymi. Jeśli jednak podwyższymy stopień wielomianów fi oraz nałożymy na nie warunki ciągłości funkcji w węzłach:

$$\bigwedge_{i=1,\ldots,K-1} f_i(\xi_i) = f_{i+1}(\xi_i) \tag{2}$$

oraz np. warunki ciągłości pochodnej rzędu pierwszego:

$$\bigwedge_{i=1,\ldots,K-1} f_i'(\xi_i) = f_{i+1}'(\xi_i) \tag{3}$$

to uzyskamy funkcję o odpowiednim stopniu gładkości.

W budowie modelu regresyjnego wykorzystującego krzywe sklejane istotnym problem jest wybranie odpowiedniego stopnia wielomianów składowych fi oraz ustalenie liczby węzłów. Proste podejście zaimplementowane

w pakiecie statystycznym „R" polega na zadaniu przez użytkownika liczby stopni swobody:

$$df = S + K \tag{4}$$

oraz stopnia wielomianów składowych. Wtedy $(df - S - 1)$ węzłów zostaje wybranych jako odpowiednie kwantyle rozkładu zmiennej X (zob. [9]).

1.1. Regularyzacja w metodzie krzywych sklejanych

Liczbę stopni swobody wyznaczamy zwykle w sposób symulacyjny, przeszukując pewien zakres parametru *df* i badając dopasowanie modelu do obserwacji ze zbioru testowego. Przyjęcie zbyt małej liczby stopni swobody prowadzi zwykle do otrzymania modelu, który daje duże błędy resubstytucji. Natomiast wybranie zbyt dużej wartości parametru *df* powoduje uzyskanie modelu nadmiernie dopasowanego do danych ze zbioru uczącego i dającego duże błędy predykcji na zbiorze testowym.

W metodzie krzywych sklejanych problem regularyzacji, czyli uzyskania kompromisu pomiędzy dopasowaniem modelu do danych ze zbioru uczącego a odpowiednim stopniem złożoności modelu, można przedstawić w postaci zadania minimalizacji funkcjonału:

$$PRSS(f,\lambda) = \sum_{j=1}^{N} (y_j - f(x_j))^2 + \lambda \int (f''(t))^2 \, dt \tag{5}$$

Pierwszy człon funkcjonału (5) jest miarą dopasowania funkcji f do obserwacji ze zbioru uczącego, natomiast drugi jest odpowiednikiem złożoności funkcji regresji. W tym przypadku złożoność modelu jest związana z gwałtownymi, oscylacyjnymi zmianami wartości estymowanej funkcji regresji. Parametr $\lambda \in (0,\infty)$, współczynnik kary, ustala proporcje pomiędzy nadmiernym dopasowaniem funkcji f a jej złożonością.

Rozwiązaniem zadania minimalizacji funkcjonału (5) jest funkcja sklejana nazywana **gładką funkcją sklejaną**:

$$\hat{f} = \min_{f} PRSS(f,\lambda) \tag{6}$$

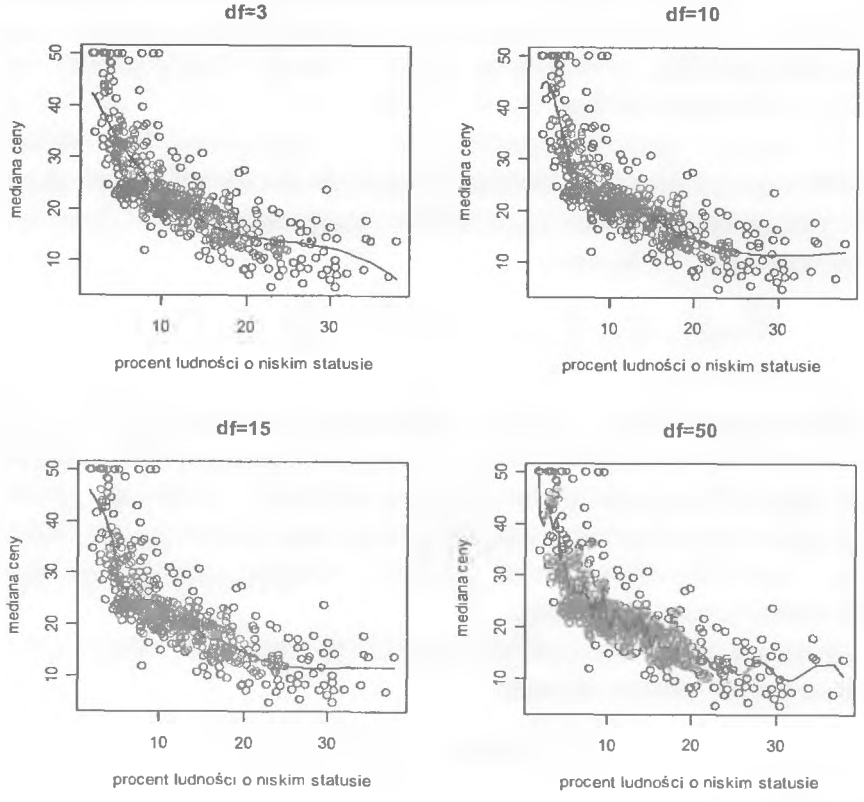
1.2. Przykład

Dla zilustrowania zagadnienia regularyzacji, na zbiorze danych *Boston*, metodą krzywych sklejanych wyznaczono modele regresji dla wybranych wartości parametru *df*.

Dane w zbiorze *Boston* zostały zebrane w 1978 roku przez Harrisona i Rubinfelda (zob. [5]), którzy zajmowali się badaniem zależności pomiędzy jakością życia a cenami nieruchomości w okolicach Bostonu.

Zbiór *Boston* jest szeroko znany i wykorzystywany do sprawdzania jakości różnych nieklasycznych metod regresji. Zawiera on 506 obserwacji, a zgromadzone dane są charakteryzowane przez 14 zmiennych (w tym jedną zmienną niemetryczną), jednak do poniższej analizy jako zmienną objaśniającą wybrano *LSTAT* – procent ludności o niskim statusie społecznym. Zmienną zależną *MEDV* jest mediana wartości domu (w tys. dolarów).

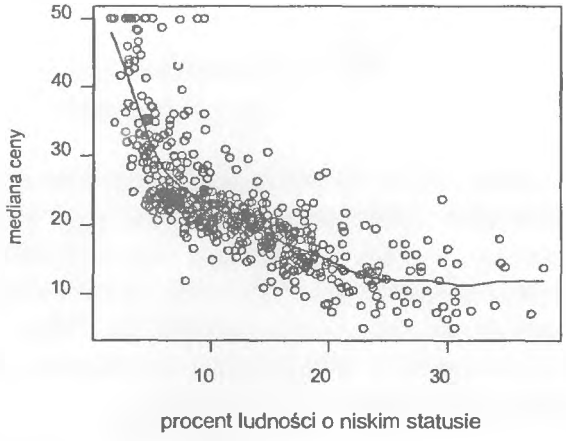
Za pomocą metody krzywych sklejanych zbudowano modele regresji dla czterech wybranych wartości parametru *df*. Otrzymane funkcje przedstawiono na rys. 1.



Rys. 1. Wykres funkcji regresji, obrazującej zależność ceny domu od procentu ludności o niskim statusie społecznym, uzyskanej za pomocą funkcji sklejanych trzeciego rzędu dla różnych stopni swobody *df*

Parametr $df = 50$ prowadzi do uzyskania funkcji sklejanej charakteryzującej się gwałtownymi, oscylacyjnymi zmianami wartości i wysokim stopniem złożoności modelu. Natomiast funkcja regresji otrzymana dla parametru $df = 3$ jest gorzej dopasowana do danych ze zbioru uczącego, ale daje niewielkie wartości miernika złożoności modelu. Wydaje się, że najbardziej optymalnym rozwiązaniem problemu regularyzacji może być model zbudowany dla $df = 10$ lub $df = 15$.

Rozwiązując zadanie minimalizacji funkcjonału (5) na zbiorze danych *Boston*, uzyskujemy gładką funkcję sklejaną dla liczby stopni swobody $df = 10$, przedstawioną na rys. 2.



Rys. 2. Wykres gładkiej funkcji sklejanej

2. Wielowymiarowa metoda krzywych sklejanych MARS

Metoda MARS (ang. multivariate adaptive regression splines) została zaproponowana przez Friedmana w 1991 roku (zob. [2]). Jest ona oparta na funkcjach sklejanych pierwszego rzędu w postaci:

$$(X-\xi)_+ = \begin{cases} X-\xi, & \text{dla } X \geq \xi \\ 0, & \text{dla } X < \xi \end{cases} \text{ oraz } (\xi - X)_+ = \begin{cases} \xi - X, & \text{dla } X \leq \xi \\ 0, & \text{dla } X > \xi \end{cases} \quad (7)$$

gdzie punkt ξ jest węzłem.

W metodzie MARS dla wszystkich zmiennych objaśniających X_j ($j = 1,\ldots, p$) tworzy się pary funkcji typu (7) z węzłami w punktach $\{x_{ij}\}_{i=1,\ldots,N}$, które są i-tymi realizacjami zmiennej X_j. W ten sposób otrzymuje się zbiór:

$$C = \left\{(X_j - \xi_j)_+, (\xi_j - X_j)_+ : \ \xi_j \in \{x_{1j}, x_{2j}, \ldots, x_{Nj}\}, \ j = 1,2,\ldots, p\right\} \quad (8)$$

Model regresyjny rozważany w omawianej metodzie jest modelem addytywnym w postaci:

$$f(\mathbf{X}) = \alpha_0 + \sum_{m=1}^{M} \alpha_m h_m(\mathbf{X}) \quad (9)$$

gdzie:

$$h_m(\mathbf{X}) = \prod_{k=1}^{d} u_m (X_k - \xi_k)_+ \quad (10)$$

dla $u_m \in \{-1,1\}$ oraz $(X_k - \xi_k)_+ \in C$.

Funkcje h_m w modelu (9) są iloczynami tensorowymi funkcji ze zbioru C, zaś wartość parametru d oznacza przyjętą liczbę interakcji pomiędzy zmiennymi X_j.

Algorytm budowy modelu regresyjnego metodą MARS składa się z dwóch głównych etapów: doboru zmiennych oraz ich eliminacji, a jego kolejne kroki zostały szczegółowo przedstawione w pracy [8]. Niniejszy artykuł skupia się jedynie na omówieniu tych kroków algorytmu, które są istotne z punktu widzenia zadania regularyzacji modelu.

2.1. Regularyzacja w metodzie MARS

Tak jak już wspomniano, algorytm metody MARS składa się z dwóch etapów: doboru zmiennych do modelu oraz ich eliminacji.

W etapie doboru zmiennych, do modelu wprowadzono funkcje ze zbioru C, dające najlepsze dopasowanie funkcji regresji do danych ze zbioru uczącego. W każdym kolejnym kroku są wybierane dwie funkcje ze zbioru C, które w największym stopniu redukują sumę kwadratów reszt i zachowują postać iloczynu tensorowego funkcji $h_m(\mathrm{X})$. Procedura ta jest powtarzana dopóki maleje błąd resubstytucji lub nie zostanie osiągnięta zadana, maksymalna liczba funkcji bazowych modelu.

W efekcie wykonania procedury doboru zmiennych otrzymujemy model, który charakteryzuje się nadmiernym dopasowaniem do danych ze zbioru uczącego oraz wysoką złożonością. Następnym etapem jest więc procedura eliminacji. Z modelu zostają usunięte te wyrażenia, których eliminacja poprawi jakość modelu lub nieco ją obniży, ale zmniejszy się złożoność modelu.

W każdym kroku z modelu zostaje wyeliminowana jedna funkcja bazowa, ta której usunięcie powoduje najmniejszy wzrost sumy kwadratów reszt. Model taki zapamiętujemy jako $\hat{f}_m$, gdzie m jest liczbą funkcji bazowych. Z sekwencji modeli zostaje wybrany ten, który jest najlepszy w sensie przyjętego kryterium.

Do oceny jakości modelu stosuje się **uogólnione kryterium sprawdzania krzyżowego** (ang. generalized cross-validation criterion):

$$\mathrm{GCV}(m) = \frac{\sum_{i=1}^{N} \left(y_i - \hat{f}_m(\mathbf{X}_i)\right)^2}{N\left(1 - \frac{Z(m)}{N}\right)^2} \tag{11}$$

Funkcja $Z(m)$ jest miarą złożoności modelu regresyjnego, który zawiera m funkcji bazowych. Proponuje się przyjęcie jej wartości na poziomie:

$$Z(m) = (1 + \gamma) \cdot \mathrm{m} \tag{12}$$

gdzie γ zmienia się od 2 do 4 i oznacza liczbę parametrów modelu związaną z funkcjami bazowymi.

Kryterium GCV(m) jest miarą wyrażającą kompromis pomiędzy dopasowaniem modelu a jego złożonością. Licznik formuły zapisanej wzorem (11) jest sumą kwadratów reszt, natomiast mianownik wyraża stopień złożoności modelu. Funkcja $\hat{f}_m$, dla której wartość uogólnionego kryterium sprawdzania krzyżowego jest największa, jest tym samym rozwiązaniem zagadnienia regularyzacji.

2.2. Przykład

Do skonstruowania przykładu ilustrującego metodę MARS ponownie wykorzystano dane ze zbioru *Boston*. Rolę zmiennej objaśnianej Y będzie nadal pełniła mediana wartości domu, natomiast jako zmienne objaśniające przyjmiemy:

$X_1 = INDUS$ – wskaźnik industrializacji,
$X_2 = NOX$ – koncentrację tlenku azotu,
$X_3 = RM$ – średnią liczbę pokoi,
$X_4 = RAD$ – dostęp do autostrady,
$X_5 = TAX$ – wysokość płaconych podatków,
$X_6 = P/T$ – liczbę uczniów przypadających na jednego nauczyciela,
$X_7 = LSTAT$ – procent ludności o niskim statusie społecznym.

Model regresyjny został zbudowany za pomocą funkcji *polymars*, zaimplementowanej w pakiecie statystycznym „R". W efekcie wywołania tej funkcji otrzymano macierz przedstawioną w tabeli 1.

Tabela 1

Kolejne kroki procedury MARS

Lp.	0/1	SIZE	RSS	GCV	Lp.	0/1	SIZE	RSS	GCV
1	1	1	42716,295	85,77	49	0	47	3390,943	16,97
2	1	2	19472,381	39,73	50	0	46	3401,125	16,60
3	1	3	14026,451	29,08	51	0	45	3427,719	16,32
4	1	4	12364,693	26,06	52	0	44	3439,031	15,98
5	1	5	11146,495	23,88	53	0	43	3458,099	15,69
6	1	6	8980,562	19,56	54	0	42	3475,683	15,40
7	1	7	8609,191	19,07	55	0	41	3505,645	15,17
8	1	8	8200,574	18,47	56	0	40	3523,225	14,90
9	1	9	7953,901	18,22	57	0	39	3554,490	14,68
10	1	10	7826,292	18,24	58	0	38	3584,810	14,47
11	1	11	7724,176	18,31	59	0	37	3609,203	14,25
12	1	12	7637,828	18,42	60	0	36	3645,717	14,08
13	1	13	7555,837	18,55	61	0	35	3694,523	13,96
14	1	14	7329,299	18,31	62	0	34	3755,735	13,88
15	1	15	7002,415	17,81	63	0	33	3769,634	13,64
16	1	16	6832,279	17,70	64	0	32	3841,140	13,60
17	1	17	6661,521	17,57	65	0	31	3853,790	13,36
18	1	18	6543,438	17,58	66	0	30	3988,827	13,55
19	1	19	6511,103	17,82	67	0	29	3997,878	13,30
20	1	20	6208,822	17,31	68	0	28	4007,145	13,06
21	1	21	5977,121	16,98	69	0	27	4133,932	13,21
22	1	22	5765,858	16,70	70	0	26	4175,193	13,07
23	1	23	5557,936	16,41	71	0	25	4304,985	13,22
24	1	24	5362,701	16,14	**72**	**0**	**24**	**4312,747**	**12,99**
25	1	25	5265,732	16,16	73	0	23	4454,835	13,15
26	1	26	5201,091	16,29	74	0	22	4598,315	13,32
27	1	27	5100,665	16,29	75	0	21	4825,679	13,71
28	1	28	5055,421	16,48	76	0	20	4917,750	13,71
29	1	29	4978,406	16,56	77	0	19	5102,429	13,96
30	1	30	4936,043	16,76	78	0	18	5335,921	14,33
31	1	31	4901,798	17,00	79	0	17	5551,673	14,64
32	1	32	4818,243	17,06	80	0	16	5803,076	15,03
33	1	33	4778,822	17,29	81	0	15	5855,902	14,90
34	1	34	4493,612	16,61	82	0	14	6208,032	15,51
35	1	35	4439,058	16,77	83	0	13	6928,527	17,01
36	1	36	4385,253	16,93	84	0	12	7744,486	18,68
37	1	37	4303,180	16,99	85	0	11	7901,370	18,73
38	1	38	4259,982	17,20	86	0	10	8079,243	18,83
39	1	39	3840,684	15,86	87	0	9	8421,305	19,29
40	1	40	3791,594	16,03	88	0	8	8875,691	19,99
41	1	41	3746,516	16,21	89	0	7	10022,678	22,20
42	1	42	3699,184	16,38	90	0	6	11528,461	25,11
43	1	43	3655,825	16,58	91	0	5	11548,646	24,74
44	1	44	3614,801	16,80	92	0	4	14901,228	31,40
45	1	45	3576,441	17,03	93	0	3	18771,081	38,92
46	1	46	3516,959	17,16	94	0	2	22061,879	45,01
47	1	47	3447,157	17,25	95	0	1	42716,295	85,77
48	1	48	3381,489	17,35					

Wiersze tej macierzy zawierają kolejne kroki procedury tworzącej model regresyjny. W drugiej kolumnie występują wartości 0 lub 1, które wskazują, z jakim etapem ma się do czynienia. Wartość 1 informuje, że następuje dodanie nowej zmiennej do modelu, natomiast 0 mówi o usunięciu wyrażenia z modelu. Kolumna *RSS* zawiera sumy kwadratów reszt liczone w kolejnych krokach procedury. Kolumna *GCV* to kolejne wartości uogólnionego kryterium sprawdzania krzyżowego. W naszym przypadku najmniejsza wartość *GCV* to 12,99 dla modelu zawierającego 24 funkcje bazowe. Jest to więc model będący rozwiązaniem zadania regularyzacji.

3. Addytywna metoda drzew regresyjnych MART

Addytywna metoda drzew regresyjnych MART (ang. multiple additive regression trees) została zaproponowana przez Friedmana w 1999 roku (zob. [3]). Należy ona do grupy metod agregacyjnych wykorzystujących sekwencyjne łączenie modeli składowych (ang. boosting).

Model regresyjny w metodzie MART można przedstawić w postaci:

$$f(\mathbf{x}) = \sum_{m=0}^{M} T(\mathbf{x}, \Theta_m) \tag{13}$$

gdzie funkcje składowe $T(\mathbf{x}, \Theta_m)$ są drzewami regresyjnymi charakteryzowanymi przez zbiór parametrów $\Theta_m = \{R_{jm}, \gamma_{jm}\}_{m=0,\ldots,M,\ j=1,\ldots,J_m}$ (zob. [6]).

Estymatory parametrów Θ_m w modelu zagregowanym (13) otrzymujemy poprzez minimalizację funkcji straty $L(y,f(\mathbf{x}))$ w zbiorze uczącym U:

$$\hat{\Theta}_m = \underset{\{\Theta_m\}_{m=0,\ldots,M}}{\arg\min} \sum_{i=1}^{N} L\left(y_i, \sum_{m=0}^{M} T(\mathbf{x}_i, \Theta_m) \right) \tag{14}$$

Zagadnienie minimalizacji funkcji straty $L(y,f(\mathbf{x}))$ jest złożone obliczeniowo, dlatego do jego rozwiązania wykorzystuje się najczęściej strategię wspinaczki (zob. [6]), polegającą na wyznaczeniu w każdym kroku algorytmu rozwiązania optymalnego jedynie w sensie lokalnym. Szczegółowo omówione etapy metody MART przedstawiono w pracy [3].

W każdej iteracji algorytmu MART do modelu zostaje dołączona kolejna funkcja – drzewo regresyjne $T(\mathbf{x}, \Theta_m)$. Wszystkie funkcje składowe są również systematycznie poprawiane, przez co uzyskujemy model bardzo dobrze dopasowany do danych ze zbioru uczącego.

Im większa liczba iteracji w algorytmie MART, tym więcej funkcji składowych wchodzi do modelu regresyjnego. Uzyskujemy więc model nadmierne dopasowany do zbioru uczącego oraz charakteryzujący się dużą złożonością. Jednym ze sposobów rozwiązania zagadnienia regularyzacji jest w tym przypadku ograniczenie wartości parametru M, czyli liczby funkcji składowych.

Optymalną liczbę funkcji składowych można wyznaczyć w sposób symulacyjny, przeszukując pewien zakres parametru M i wyznaczając błędy predykcji modelu na zbiorze testowym.

3.1. Przykład

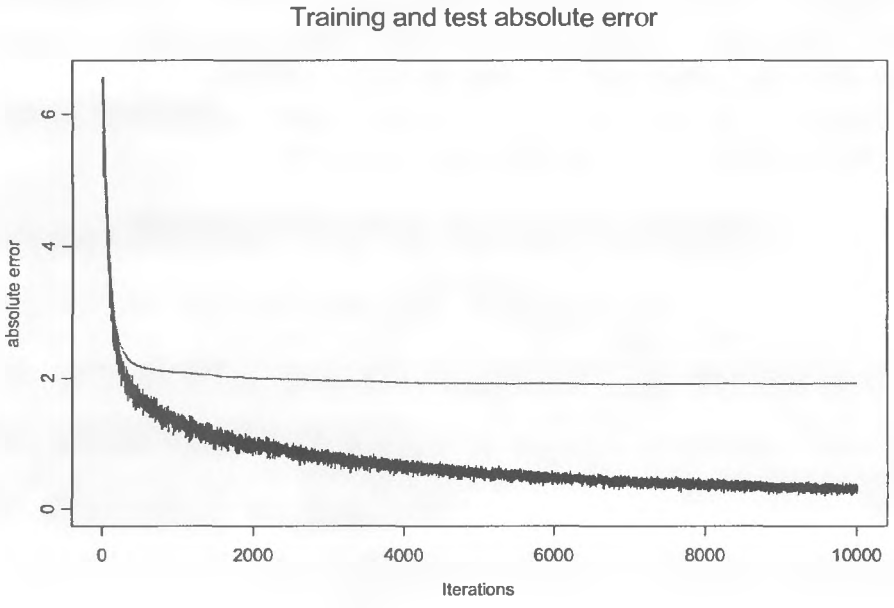
Zbiór *Boston* został losowo podzielony na zbiór uczący (zawierający 80% obserwacji) oraz zbiór testowy. Na zbiorze uczącym, za pomocą metody MART, zbudowano modele regresji o różnej liczbie funkcji składowych. Do każdego modelu wprowadzono 13 zmiennych objaśniających. Zmienną zależną jest mediana wartości domu (w tys. dolarów). Maksymalna wartość parametru M jest równa 10 000, zatem najbardziej złożony model składa się z 10 000 drzew regresyjnych.

Do oszacowania parametrów modeli wykorzystano funkcję straty Hubera

$$L(y, f(\mathbf{x})) = \begin{cases} (y - f(\mathbf{x}))^2, & \text{gdy } |y - f(\mathbf{x})| < \delta \\ \delta\left(2|y - f(\mathbf{x})| - \delta\right), & \text{gdy } |y - f(\mathbf{x})| \geq \delta \end{cases} \tag{15}$$

Poniższy wykres (rys. 3.) przedstawia zależność błędu absolutnego obliczonego na zbiorze uczącym (niższa krzywa) oraz testowym (wyższa krzywa) w zależności od liczby funkcji składowych.

Z wykresu możemy odczytać, iż błąd absolutny obliczony na zbiorze testowym stabilizuje się na poziomie 2. Zatem dokładanie do modelu kolejnych funkcji składowych i tym samym zwiększanie złożoności modelu nie doprowadzi do obniżenia wartości błędu predykcji. Rozwiązaniem zadania regularyzacji w tym przykładzie będzie model zawierający około 1000 drzew składowych.



Rys. 3. Wykres zależności błędu absolutnego obliczonego na zbiorze uczącym (niższa krzywa) oraz testowym (wyższa krzywa) w zależności od liczby funkcji składowych

Literatura

1. Cherkassky V., Mulier F.: Learning from Data – Concepts, Theory, and Methods. John Wiley & Sons, Inc., New York 1998.
2. Friedman J.H.: Multivariate Adaptive Regression Splines. „Annals of Statistics" 1991, 19, pp. 1-141.
3. Friedman J.H.: Greedy Function Approximation: A Gradient Boosting Machine. Technical report, Dept. of Statistics, Stanford University, 1999.
4. Friedman J.H.: Stochastic Gradient Boosting. Technical report, Dept. of Statistics, Stanford University, 1999.
5. Harrison D., Rubinfeld D.L.: Hedonic Prices and the Demand for Clean Air. „Journal of Environmental Economics and Management 8" 1978.
6. Hastie T., Tibshirani R., Friedman J.H.: The Elements of Statistical Learning. Springer-Verlag, New York 2001.
7. Koronacki J., Ćwik J.: Statystyczne systemy uczące się. Wydawnictwo Naukowo-Techniczne, Warszawa 2005.

8. Trzęsiok J.: Wybrane nieparametryczne metody regresji i ich zastosowania. W: Klasyfikacja i analiza danych – teoria i zastosowania. Red. K. Jajuga, M. Walesiak. „Taksonomia 11" 2004, nr 1022, s. 107-115.
9. Trzęsiok J.: Metoda krzywych składanych w budowie modelu regresyjnego. Zeszyty Naukowe, nr 36, Katowice 2005, s. 171-182.

VARIOUS REGULARIZATION ISSUES OF REGRESSION

Summary

It is well known in statistics, that fitting the training data too well can increase prediction risk on the future predictions. In other words too large flexibility of the regression function would cause a learner to overfit the data, i.e. the learner would be able to model the noise in the data as well as the generating process and it leads to poor generalization. The process of finding the balance between minimizing the training error and controlling capacity is called regularization. The paper presents the issue and gives two examples of regularization technique in case of MARS and MART.

Michał Trzęsiok

EMPIRYCZNA OCENA WRAŻLIWOŚCI METODY WEKTORÓW NOŚNYCH NA WYSTĘPOWANIE OBIEKTÓW BŁĘDNIE SKLASYFIKOWANYCH W ZBIORZE UCZĄCYM

Wprowadzenie

Empiryczne badania porównawcze wskazują, że metoda wektorów nośnych (ang. Support Vector Machines – SVM) jest, obok metody zagregowanych drzew klasyfikacyjnych Breimana, najefektywniejszą metodą dyskryminacji [4]. Przez efektywność rozumiemy tu przede wszystkim małe błędy klasyfikacji dla obiektów ze zbioru rozpoznawanego, ale także możliwość stosowania metody dla różnych typów danych.

Funkcja klasyfikująca, otrzymywana metodą wektorów nośnych, jest nieliniowa i wybierana z przestrzeni hipotez zawierającej wiele bardzo zróżnicowanych funkcji. Powoduje to zagrożenie wystąpienia zjawiska nadmiernego dopasowania funkcji klasyfikującej do danych ze zbioru uczącego (ang. overfitting), na których model został zbudowany. Otrzymany model będzie wtedy opisywał badane zjawisko zbyt szczegółowo, rygorystycznie, opierając się na zbiorze uczącym, w którym, jak wiadomo, wartości cech diagnostycznych są realizacjami zmiennych losowych i są obarczone błędami. Problem ten został rozwiązany przez zaimplementowanie w algorytmie metody wektorów nośnych zasady minimalizacji ryzyka strukturalnego [6; 5]. Najogólniej rzecz biorąc, zasada minimalizacji ryzyka strukturalnego polega na poszukiwaniu funkcji dyskryminującej jak najlepiej klasyfikującej obserwacje ze zbioru uczącego, jednocześnie na kontrolowaniu stopnia złożoności wyznaczanej funkcji.

W rzeczywistych zbiorach danych podlegających dyskryminacji bardzo często występują obiekty błędnie sklasyfikowane. W takim przypadku wymaganie zerowego lub bliskiego zeru błędu klasyfikacji na zbiorze uczącym jest niezasadne, gdyż osłabia zdolność modelu do poprawnego klasyfikowania nowych obiektów ze zbioru rozpoznawanego. Wszystkie metody, które za podstawowe i jedyne kryterium wyboru funkcji dyskryminującej biorą minimalizowanie błędu klasyfikacji na zbiorze uczącym, są więc uważane za mało elastyczne i nieodporne.

Metoda wektorów nośnych jest uważana za metodę odporną. W dalszej części artykułu przedstawiono pokrótce algorytm metody SVM, ze szczególnym uwzględnieniem elementów czyniących ją odporną na błędy występujące w zbiorze uczącym, a następnie empirycznie sprawdzono na zbiorze danych standardowo wykorzystywanym do badania własności metod wielowymiarowej analizy statystycznej, w jakim stopniu metoda jest odporna. Dla porównania zbadano również konkurencyjne metody dyskryminacji.

1. Metoda wektorów nośnych

Dany jest zbiór uczący $D = \{(\mathbf{x}^1, y^1), \ldots, (\mathbf{x}^N, y^N)\}$, gdzie $\mathbf{x}^i \in \mathbf{R}^d$ oraz y^i – wartości zmiennej opisującej klasę obiektu. Metoda wektorów nośnych, realizując nieliniową klasyfikację, w pierwszej kolejności transformuje obserwacje z oryginalnej przestrzeni danych w przestrzeń o dużo większym wymiarze, w której obiekty są rozdzielane hiperpłaszczyznami. Ze względu na nieliniowość przekształcenia przestrzeni danych, liniowemu rozdzieleniu danych w nowej przestrzeni cech odpowiada nieliniowa ich dyskryminacja w przestrzeni pierwotnej.

Jeżeli przez φ oznaczymy nieliniową transformację przestrzeni danych, to w przypadku dwóch klas zadanie dyskryminacji polega na wyznaczeniu optymalnej hiperpłaszczyzny:

$$\boldsymbol{\beta} \cdot \varphi(\mathbf{x}) + \beta_0 = 0 \tag{1}$$

rozdzielającej klasy zbioru uczącego $\{(\varphi(\mathbf{x}^1), y^1), \ldots, (\varphi(\mathbf{x}^N), y^N)\}$, gdzie $\mathbf{x}^i \in R^d$, $\varphi(\mathbf{x}^i) \in Z$ oraz $y^i \in \{-1,1\}$ dla $i = 1,\ldots,N$. W przypadku większej liczby klas można wyznaczyć wiele funkcji dyskryminujących klasy parami i skonstruować funkcję dyskryminującą, korzystając z reguły majoryzacji (głosowania modeli cząstkowych).

Jak pokazano m.in. w pracy [1], rozważane zagadnienie można zapisać w postaci zadania optymalizacji wypukłej z kwadratową funkcją celu oraz liniowymi ograniczeniami:

$$\begin{cases} \min\limits_{\beta,\beta_0} \frac{1}{2}\|\beta\|^2 + C\sum\limits_{i=1}^{N}\xi_i, \\ \xi_i \geq 0, \quad y^i\left(\beta\cdot\varphi(\mathbf{x}^i)+\beta_0\right)\geq 1-\xi_i, \quad i=1,\ldots,N. \end{cases} \tag{2}$$

Nierównościowe ograniczenia w (1) w postaci $y^i\left(\beta\cdot\varphi(\mathbf{x}^i)+\beta_0\right)\geq 1$ z geometrycznego punktu widzenia stanowią warunki separowalności, tzn. wymuszają, aby wyznaczona hiperpłaszczyzna rozdzielała klasy. Wprowadzone zaś zmienne $\xi_1,\ldots\ \xi_N \geq 0$ są realizacją postulatu uelastycznienia metody, gdyż osłabiają wymaganie, aby wszystkie obserwacje ze zbioru uczącego były poprawnie klasyfikowane (rozdzielone) przez hiperpłaszczyznę.

Rozwiązanie zadania (2) znajdujemy metodą mnożników Lagrange'a. Funkcję dyskryminującą otrzymujemy wykorzystując formułę definiującą optymalną hiperpłaszczyznę rozdzielającą klasy:

$$f(\mathbf{x}) = \operatorname{sign}\left[\sum_{i\in I_{SV}} \alpha_i y^i \varphi(\mathbf{x}^i)\cdot\varphi(\mathbf{x}) + \hat{\beta}_0\right] \tag{3}$$

i jest ona opisana wyłącznie przez te wektory $\mathbf{x}^i$ ze zbioru uczącego, którym odpowiadają niezerowe współczynniki Lagrange'a w rozwiązaniu zadania optymalizacyjnego (2). Obserwacje te nazywamy wektorami nośnymi. Ponadto w metodzie wektorów nośnych wykorzystuje się funkcje z rodziny funkcji jądrowych, definiujące iloczyn skalarny w pewnej przestrzeni cech. Tym samym postać funkcji transformującej φ nie musi być znana. Wystarczy bowiem postać iloczynu skalarnego $K(u,v) = \varphi(u)\cdot\varphi(v)$ w przestrzeni Z. Własność ta ma istotne znaczenie dla wyeliminowania wielu numerycznych problemów metody [3].

2. Analiza wrażliwości metod na występowanie obserwacji błędnie sklasyfikowanych w zbiorze uczącym

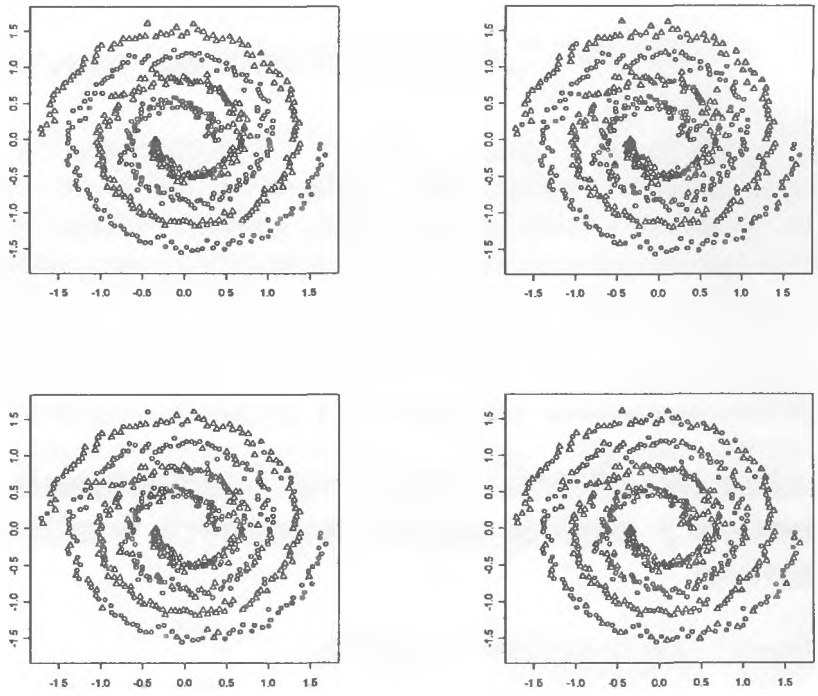
2.1. Zbiory danych poddane analizie

W dalszych rozważaniach przedstawiono porównanie jakości modeli otrzymanych za pomocą różnych metod dyskryminacji w przypadku zbiorów uczących, w których z góry zadana część obiektów została błędnie sklasyfikowana. Tak postawiony problem w zasadzie wyklucza możliwość wykorzystania zbiorów danych rzeczywistych, gdyż w ich przypadku nie jest znana poprawna przynależność do klas badanych obiektów. Do analizy wykorzystano zbiór da-

nych *Spirals* – zbiór sztuczny, wygenerowany za pomocą funkcji z biblioteki mlbench pakietu statystycznego R [2]. Zbiór *Spirals* tworzą punkty dwóch spiral o wspólnym początku, z których drugą otrzymano z pierwszej przez obrót o 180°, przy czym punkty obu spiral wygenerowano z uwzględnieniem zaburzenia – zmiennej o rozkładzie $N(0; 0{,}05)$. Należy zauważyć, że klasy tworzące ten zbiór są liniowo nieseparowalne, o wysokim „stopniu nieliniowości" funkcji rozdzielającej badane klasy. Zbiór uczący, zawierający 600 obiektów, został wygenerowany w czterech wariantach:

a) niezawierającym obserwacji błędnie sklasyfikowanych,
b) zawierającym 10% losowo wybranych obserwacji błędnie sklasyfikowanych,
c) zawierającym 20% losowo wybranych obserwacji błędnie sklasyfikowanych,
d) zawierającym 30% losowo wybranych obserwacji błędnie sklasyfikowanych.

Ilustrację zbiorów uczących poddanych dyskryminacji przedstawia rys. 1.



Rys. 1. Cztery warianty zbioru uczącego *Spirals*, zawierające po 600 obiektów każdy. Zbiór tworzą dwie klasy – spirale – oznaczone na rysunku kółkami oraz trójkątami. W lewym górnym rogu – zbiór, w którym wszystkie obserwacje są poprawnie sklasyfikowane. W prawym górnym rogu zbiór, w którym 10% losowo wybranych obserwacji jest błędnie sklasyfikowanych. W lewym dolnym – 20% błędnie sklasyfikowanych, zaś w prawym dolnym – aż 30%

2.2. Porównywane metody dyskryminacji oraz kryterium porównawcze

Ponieważ podstawowym celem badania jest sprawdzenie, czy metoda potrafi zbudować model, który będzie poprawnie klasyfikował nowe obiekty ze zbioru rozpoznawanego, więc jako kryterium porównywania modeli przyjęto błąd klasyfikacji liczony na zbiorze testowym, zawierającym 400 obserwacji, które nie były wykorzystywane w procesie wyznaczania funkcji dyskryminującej (modelu).

Wprawdzie nadrzędnym celem było zbadanie wrażliwości metody wektorów nośnych na występowanie obserwacji błędnie sklasyfikowanych w zbiorze uczącym, jednak, aby mieć pewien punkt odniesienia, te same zbiory poddano dyskryminacji również konkurencyjnymi metodami. W analizie uwzględniono:

1) SVM – metodę wektorów nośnych,
2) LDA – liniową analizę dyskryminacyjną,
3) KNN – metodę *k*-najbliższych sąsiadów,
4) RPART – drzewa klasyfikacyjne,
5) RFOREST – zagregowane drzewa klasyfikacyjne Breimana.

Symbole, którymi oznaczono poszczególne metody, zostały zaczerpnięte z nazw odpowiadających im funkcji z pakietu statystycznego R, który wraz z dodatkowymi bibliotekami został wykorzystany do przeprowadzenia analizy.

Prawie wszystkie badane metody wymagają ustalenia wartości pewnych parametrów. Z reguły jakość otrzymanego modelu silnie zależy od odpowiedniego doboru wartości parametrów. W przeprowadzonej procedurze badawczej wybierano taki układ wartości parametrów, przeszukując odpowiednio duży ich zakres, który dawał najmniejszy błąd klasyfikacji liczony metodą sprawdzania krzyżowego z podziałem zbioru uczącego na 10 części.

W metodzie wektorów nośnych wykorzystano funkcję jądrową Gaussa, zmieniając wartość parametru gamma w zakresie {1, 10, 20} oraz parametru C od 10^{-2} do 10^{2}. W metodzie *k*-najbliższych sąsiadów wartość parametru k dobierano z zakresu od 1 do 10. W przypadku drzew klasyfikacyjnych wymaganą minimalną liczbę obserwacji w węźle, aby nastąpił dalszy podział, ustalano na poziomie od 3 do 10, zaś kryterium minimalnej poprawy jakości modelu na poziomie od 1% do 3%. W metodzie zagregowanych drzew klasyfikacyjnych Breimana liczbę zmiennych losowanych przy każdym podziale ustalano na poziomie $\frac{\sqrt{d}}{2}\sqrt{d}$ oraz $2\sqrt{d}$, gdzie d – liczba zmiennych. Ustalono również liczbę drzew równą 100 oraz 200, a także minimalną liczbę obserwacji w liściu: 1, 5, 10.

3. Wyniki analizy

Jak już zostało wspomniane, dla każdego z czterech wariantów zbioru uczącego skonstruowano różnymi metodami modele z optymalnym, ze względu na błąd klasyfikacji liczony metodą sprawdzania krzyżowego, układem wartości parametrów. Następnie każdy z tak otrzymanych modeli wykorzystano do klasyfikacji obiektów ze zbioru testowego i obliczono błąd klasyfikacji. Wyniki analizy przedstawia tabela 1.

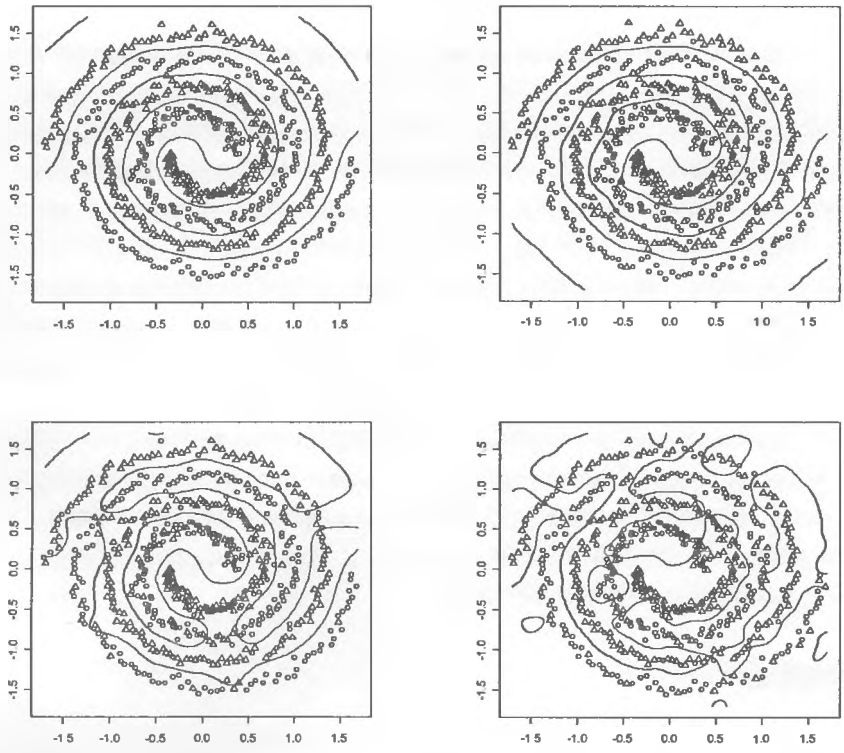
Tabela 1

Błąd klasyfikacji w % liczony na zbiorze testowym dla różnych metod dyskryminacji i zbiorów uczących *Spirals* zawierających odpowiednio 0%, 10%, 20% i 30% obserwacji błędnie sklasyfikowanych

	0% błędnych	10% błędnych	20% błędnych	30% błędnych
SVM	0,0	0,0	4,5	11,3
LDA	49,8	50,0	49,8	48,3
KNN	0,3	9,3	19,8	30,3
RPART	6,5	10,0	22,3	31,5
RFOREST	2,8	4,8	12,3	21,3

W pierwszej kolejności można zauważyć, że mimo nieliniowej postaci funkcji rozdzielających klasy, metoda wektorów nośnych wyznaczyła model, który bezbłędnie sklasyfikował obiekty ze zbioru testowego zarówno w przypadku zbioru pozbawionego obserwacji błędnie sklasyfikowanych, jak i w przypadku zbioru uczącego zawierającego ich 10%. Ponadto nawet gdy w zbiorze uczącym znajdowało się aż 20% błędnie sklasyfikowanych obiektów, metoda SVM nie poddała się zjawisku nadmiernego dopasowania do zbioru uczącego i wygenerowała reguły klasyfikacji w bardzo dużym stopniu odzwierciedlające reguły leżące u podstaw wygenerowanego zbioru, co potwierdza zaledwie 4,5% błąd klasyfikacji na zbiorze testowym. Oznacza to, że jakość predykcji modelu SVM może być o wiele wyższa niż jakość danych, które stanowiły podstawę budowy modelu. Wniosek ten potwierdził się nawet w przypadku zbioru uczącego o bardzo niskiej jakości danych, w którym aż 30% obserwacji zostało błędnie sklasyfikowanych. Model otrzymany metodą wektorów nośnych na tym zbiorze charakteryzuje się już 11,3% błędem w zbiorze testowym, lecz nadal jest to błąd znacznie niższy niż błąd samych danych zbioru uczącego.

Na rys. 2 przedstawiono dla czterech wariantów zbioru uczącego nieliniowe funkcje dyskryminujące otrzymane metodą SVM, odpowiadające liniowym hiperplaszczyznom rozdzielającym klasy w przestrzeni cech o większym wymiarze.



Rys. 2. Nieliniowe funkcje dyskryminujące dla czterech wariantów zbioru uczącego *Spirals*. Zbiór tworzą dwie klasy – spirale – oznaczone na rysunku kółkami oraz trójkątami. W lewym górnym rogu zbiór, w którym wszystkie obserwacje ze zbioru uczącego są poprawnie sklasyfikowane. W prawym górnym rogu zbiór, w którym 10% losowo wybranych obserwacji ze zbioru uczącego jest błędnie sklasyfikowanych. W lewym dolnym – 20% błędnie sklasyfikowanych, zaś w prawym dolnym – 30%

Poddając dalszej analizie wyniki zebrane w tabeli 1, można zauważyć, że w każdym przypadku metoda wektorów nośnych dała najmniejszy błąd klasyfikacji na zbiorze testowym. Potwierdza to wyniki otrzymane wcześniej w pracy [4]. Nie są zaskakujące duże błędy klasyfikacji dla liniowej analizy dyskryminacyjnej ze względu na ewidentną nieliniowość postaci reguł klasyfikacyjnych.

Metoda *k*-najbliższych sąsiadów, której algorytm określa przynależność danego obiektu do klasy lokalnie, biorąc pod uwagę tylko ustaloną z góry liczbę obserwacji najbliższych, sprawdzała się jedynie w przypadku zbiorów separowalnych (niekoniecznie liniowo separowalnych), tzn. algorytm dobrze radzi sobie z nieliniowością, lecz wyniki znacznie pogarszają się w przypadku, gdy klasy częściowo się pokrywają.

Podsumowanie

Wyniki badania empirycznego potwierdziły, że metoda wektorów nośnych należy do grupy metod odpornych w tym sensie, że dopuszcza, aby w zbiorze uczącym znajdowały się obserwacje błędnie sklasyfikowane. Nawet przy stosunkowo dużej liczbie takich obserwacji metoda SVM buduje reguły poprawnie klasyfikujące obiekty ze zbioru rozpoznawanego. Bardzo często w przypadku metod dopuszczających złożone, nieliniowe postaci funkcji dyskryminujących występuje problem nadmiernego dopasowania modelu do zbioru uczącego. Metoda SVM poprzez zaimplementowaną zasadę minimalizacji ryzyka strukturalnego umożliwia nieliniową dyskryminację, zachowując zdolność do poprawnego klasyfikowania obserwacji ze zbioru rozpoznawanego.

Porównanie błędów klasyfikacji na badanym zbiorze wskazuje, że metoda SVM najlepiej spośród porównywanych metod dyskryminacji nadaje się do klasyfikacji zbiorów obarczonych błędami. Niezależnie też od wariantu zbioru, metoda wektorów nośnych wyznaczała funkcję dyskryminującą, charakteryzującą się najmniejszymi błędami klasyfikacji.

Literatura

1. Cristianini N., Shawe-Taylor J.: An Introduction To Support Vector Machines (and Other Kernel-Based Learning Methods). Cambridge University Press, Cambridge 2000.
2. Leisch F., Dimitriadou E.: The mlbench Package – A Collection for Artificial and Real-World Machine Learning Benchmarking Problems. R Package, Version 1.0-0. Dostępne przez: http://cran.R-project.org (2004).
3. Smola A., Schölkopf B.: Learning with Kernels. Support Vector Machines, Regularization, Optimization, and Beyond. MIT Press, Cambridge 2002.
4. Trzęsiok M.: Analiza wybranych własności metody dyskryminacji wykorzystującej wektory nośne. W: Postępy ekonometrii. Red. A.S. Barczak. AE, Katowice 2004.
5. Trzęsiok M.: Zarys teoretycznych podstaw metody dyskryminacji wykorzystującej wektory nośne. „Studia Ekonomiczne" 2005, nr 36.
6. Vapnik V: Statistical Learning Theory. John Wiley & Sons, N.Y. 1998.

SUPPORT VECTOR CLASSIFICATION AND ITS SENSITIVITY TO THE PRESENCE OF NOISE IN THE TRAINING DATA

Summary

The Support Vector Machines have been developed as a robust tool for classification in noisy, complex domains. The paper presents a comparison of some selected classification methods by the means of classification test set error depending on the presence of noise in the training data.

Henryk Zawadzki

OBLICZENIOWE ASPEKTY WYZNACZANIA WEWNĘTRZNEJ STOPY ZWROTU DLA INWESTYCJI Z GĘSTYMI NOŚNIKAMI PRZEPŁYWÓW FINANSOWYCH

Wstęp

Wewnętrzna stopa zwrotu inwestycji (*IRR*), mimo wad, nadal jest jednym z podstawowych kryteriów stosowanych w finansowych (dynamicznych) metodach oceny projektów inwestycyjnych*. W przypadku gdy przepływy finansowe mają miejsce w skończonej liczbie momentów czasowych (dla uproszczenia załóżmy, że momentami tymi są $t = 0, 1, 2, \ldots, n$), wyznaczenie *IRR* sprowadza się do znalezienia dodatniego pierwiastka r równania:

$$NPV(r) = \sum_{t=0}^{n} \frac{CIF_t}{(1+r)^t} + \sum_{t=0}^{n} \frac{COF_t}{(1+r)^t} = 0 \qquad (1)$$

w którym CIF_t i COF_t oznaczają odpowiednio wpływy i wypływy pieniężne (przychody lub wydatki) w momencie t, natomiast n oznacza zakładany okres eksploatacji inwestycji. Rozwiązywanie równań postaci (1) stało się proste, gdy pojawiły się programy zwane systemami algebry komputerowej (Computer Algebra Systems, CAS), m.in. takie jak *Derive*, *Maple*, *Mathcad*, *Matlab*, czy *Mathematica*. Za pomocą tych programów można bez trudu znaleźć dokładne lub przybliżone wartości wszystkich pierwiastków tych równań, nawet dla dużych wartości n [9; 10].

* Zob. np. [2; 4, 7].

Gdy mamy do czynienia z inwestycjami o gęstych nośnikach przepływów finansowych, a dokładniej, gdy zbiór momentów czasowych, w których mają miejsce niezerowe przepływy finansowe, jest zbiorem w sobie gęstym*, stopa zwrotu *IRR* z inwestycji *X* jest równa:

$$IRR = 1/(1+\xi) \tag{2}$$

gdzie ξ oznacza jedyny pierwiastek równania:

$$\int_0^T cif_X(t)\xi^t dt = -\int_0^T cof_X(t)\xi^t dt \tag{3}$$

należący do przedziału (0, 1)**. W powyższym wzorze $T > 0$ jest okresem eksploatacji inwestycji (projektu), a funkcje $cif_X : [0, T] \to R^+$ oraz $cof_X : [0, T] \to R^-$ oznaczają odpowiednio intensywność przychodów i wydatków***.

Okazuje się, że nawet w prostych przypadkach, całki występujące w równaniu (3) mogą być całkami nieelementarnymi, a samo równanie równaniem przestępnym, niekiedy tak złożonym, że jego rozwiązanie jest możliwe wyłącznie za pomocą wyżej wspomnianych programów.

W artykule rozważono przypadki, w których intensywności przepływów pieniężnych są funkcjami wykładniczymi i potęgowymi oraz przedstawiono otrzymane za pomocą programu *Mathematica*® rozwiązania równania (3) oraz wartości *IRR* dla obu tych przypadków.

1. Równania przestępne a *Mathematica*®

Niech funkcja $y = f(x)$ ($f : R \to R$) będzie określona w pewnym otwartym zbiorze $D \subseteq R$. Funkcję f nazywamy analityczną, jeżeli każdy punkt $x_0 \in D$ ma otoczenie $U(x_0) \subset D$ takie, że f w tym otoczeniu daje się przedstawić w postaci szeregu potęgowego o środku x_0. Funkcję f nazywamy funkcją algebraiczną zmiennej x wtedy i tylko wtedy, gdy x i y spełniają równanie uwikłane postaci $F(x, y) = 0$, w którym F jest wielomianem zmiennych x i y. Funkcją przestępną nazywamy każdą funkcję analityczną, która nie jest algebraiczna. Najprostszymi

* Zbiorem w sobie gęstym nazywamy zbiór $A \subset X$ (X – przestrzeń topologiczna), którego każdy punkt jest jego punktem skupienia [3]. W artykule rozważa się wyłącznie przypadek, gdy nośniki przepływów finansowych są przedziałami domkniętymi.

** Zob. [5].

*** R^+ i R^- oznaczają odpowiednio zbiór liczb rzeczywistych nieujemnych i zbiór liczb rzeczywistych niedodatnich.

przykładami funkcji przestępnych są funkcje trygonometryczne, funkcja wykładnicza i funkcja logarytmiczna.

Równaniami przestępnymi nazywamy równania, w których występują funkcje przestępne. Przykładami równań przestępnych są więc równania trygonometryczne, wykładnicze, logarytmiczne oraz m.in. takie równania, jak:

$$\sin x + 2\ln x - x = 0 \tag{4}$$

Ogólne metody rozwiązywania równań przestępnych są znane jedynie dla niektórych typów tych równań. Równanie (4) oraz równania przestępne występujące w dalszej części artykułu można rozwiązać jedynie w sposób przybliżony, wykorzystując metodę połowienia przedziału, metodę siecznych, metodę Newtona (zwaną również metodą stycznych), metodę punktu stałego lub też ich modyfikacje [6].

Mathematica® jest programem komputerowym, umożliwiającym wykonywanie złożonych obliczeń numerycznych i symbolicznych oraz tworzenie dwu- i trójwymiarowej grafiki. Za tym lakonicznym określeniem kryje się program albo, jak piszą niektórzy, „zintegrowane środowisko" o olbrzymich możliwościach, które użytkownik może jeszcze poszerzyć i dostosować do własnych potrzeb, pisząc własne aplikacje w oferowanym przez program *Mathematica*® języku programowania wysokiego poziomu. Najnowsze wersje programu mogą ponadto pełnić rolę edytora tekstów (w tym naukowych, a w szczególności matematycznych) oraz umożliwiają m.in. eksport utworzonych dokumentów w formatach HTML, TEX oraz LATEX. Twórcą programu *Mathematica*®, a dokładniej jego pierwszej wersji, która ukazała się 23 kwietnia 1988 roku, jest Stephen Wolfram. Kolejne, udoskonalone wersje programu, pisane już przy współudziale coraz liczniejszego grona programistów oraz naukowców różnych specjalności i sprzedawane jako produkt firmy Wolfram Research Inc. z siedzibą w Champaign, IL, USA, stały się światowym standardem w dziedzinie obliczeń naukowych. Najnowszą wersją programu jest obecnie *Mathematica*® 6.0. Obszerną informację na jego temat można znaleźć w pozycjach [1; 8] oraz w Internecie na stronie **www.wolfram.com**.

W programie *Mathematica*® za pomocą instrukcji:

```
FindRoot[f[x]==0,{x, x0}]
```

wykorzystującej metodę Newtona, można wyznaczyć numeryczne rozwiązanie równania $f(x) = 0$, biorąc punkt x_0 jako pierwsze przybliżenie. W doborze punk-

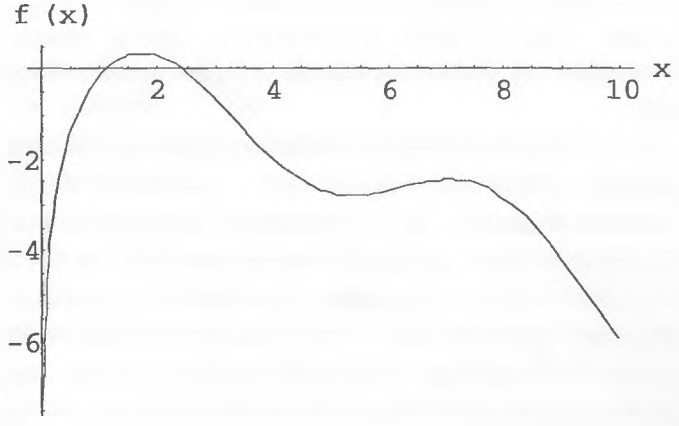
tu x_0, czyli wstępnej lokalizacji pierwiastka równania, pomaga narysowanie wykresu funkcji f, który można sporządzić za pomocą instrukcji:

```
Plot[f[x],{x, a, b}]
```

Instrukcja ta (mająca wiele wariantów i opcji) rysuje wykres funkcji f w przedziale $[a, b]$. Przykładowo, dla równania (4) instrukcja:

```
Plot[Sin[x]+2*Log[x]-x,{x,0,10},AxesLabel→{"x","f(x)"}]
```

rysuje wykres funkcji stanowiącej lewą stronę równania (4) dla $x \in (0, 10]$ (rys. 1).



Rys. 1. Wykres funkcji $f(x) = \sin x + 2 \ln x - x$ w przedziale (0, 10]

Na rys. 1 widać, że pierwiastki równania (4) znajdują się w pobliżu liczb 1 i 2. Instrukcje:

```
FindRoot[Sin[x]+2*Log[x]-x==0,{x,1}]
```

oraz:

```
FindRoot[Sin[x]+2*Log[x]-x==0,{x,2}]
```

wyznaczają przybliżone wartości jedynych pierwiastków $x \approx 1.1145$ oraz $x \approx$ ≈ 2.42874 równania (4). Stosując odpowiednie opcje, możemy zwiększyć dokładność otrzymanych wyników.

Wszystkie obliczenia (całkowanie, rozwiązywanie równań przestępnych) oraz rysunki w tym artykule zostały wykonane przez Autora za pomocą programu *Mathematica*® (wersja 5.2).

2. IRR jako pierwiastek równania przestępnego

Rozważmy najpierw przypadek, w którym intensywności przychodów i wydatków są funkcjami wykładniczymi:

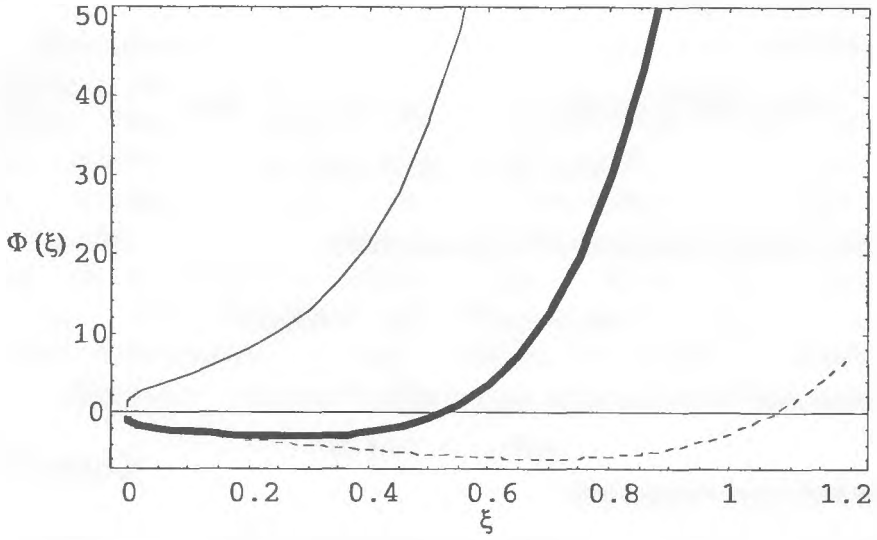
$$cif_X(t) = \alpha e^{\beta t}, \quad cof_X(t) = \gamma e^{-\delta t} \tag{5}$$

gdzie $\alpha, \beta, \gamma, \delta > 0$. Całki występujące po obu stronach równania (3) są wtedy łatwymi do obliczenia całkami elementarnymi, natomiast równanie (3) jest ze względu na ξ równaniem przestępnym w postaci:

$$\Phi(\xi) = \frac{\alpha\left(e^{(\beta+\ln\xi)T} - 1\right)}{\beta + \ln\xi} + \frac{\gamma\left(1 - e^{(-\delta+\ln\xi)T}\right)}{-\delta + \ln\xi} = 0 \tag{6}$$

Na istnienie i położenie pierwiastka ξ tego równania w przedziale $(0, \infty)$ ma istotny wpływ każdy z parametrów $\alpha, \beta, \gamma, \delta$ i T.

Dla wartości $\alpha = 2$, $\beta = 0.5$, $\gamma = 10$, $\delta = 0.5$ i $T = 7$ [5] równanie (6) ma pierwiastek $\xi \approx 0.522797$, dla którego z równania (2) otrzymujemy $IRR \approx$ ≈ 0.912788. Na rys. 2 pokazano wykres funkcji Φ (lewej strony równania (6)) dla podanych wyżej wartości parametrów (linia ciągła pogrubiona).



Rys. 2. Wykresy funkcji Φ dla różnych wartości parametrów α, β, γ i δ

Gdy zmniejszymy wartość parametru β do 0.05 (ceteris paribus), równanie (6) nie będzie miało pierwiastka w przedziale (0, 1). Pierwiastkiem jest wte-

dy $\xi \approx 1.07661$, co daje $IRR \approx -0.0711562$. Wykres funkcji Φ z równania (6) pokazano na rys. 2 (linia przerywana).

W przypadku gdy parametry przyjmą wartości $\alpha = 10$, $\beta = 0.5$, $\gamma = 2$, $\delta = 0.5$ i $T = 7$, równanie przestępne (6) w ogóle nie będzie miało dodatniego pierwiastka (rys. 2, linia ciągła).

Niech teraz intensywności przychodów i wydatków będą funkcjami potęgowymi w postaci:

$$cif_X(t) = \sqrt{t} + 1, \quad cof_X(t) = \sqrt[3]{t} - 3$$

a czas życia ekonomicznego projektu wynosi $T = 5$. W tym przypadku każda z całek występujących w równaniu (3) jest całką nie elementarną, a równanie (3) jest równaniem przestępnym:

$$\Phi(\xi) = \frac{(1+\sqrt{5})\xi^5 - 1}{\ln \xi} - \frac{1}{3(-\ln \xi)^{4/3}}\left[3\xi^5(\sqrt[3]{5} - 3)\sqrt[3]{-\ln \xi} + \Gamma\left(\frac{1}{3}, -5\ln \xi\right) + 9\sqrt[3]{-\ln \xi} - 3\Gamma\left(\frac{4}{3}\right)\right] - \frac{\sqrt{\pi}\,erfi(\sqrt{5}\ln^{\frac{1}{2}} \xi)}{2\ln^{\frac{3}{2}} \xi} = 0 \tag{7}$$

W powyższym równaniu $\Gamma(z)$ jest funkcją gamma Eulera:

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt, \quad \mathrm{Re}(z) > 0$$

funkcja $\Gamma(a,z)$ jest niepełną funkcją gamma Eulera:

$$\Gamma(a,z) = \int_z^\infty e^{-t} t^{a-1} dt, \quad \mathrm{Re}(a) > 0$$

a funkcja erfi (z) jest zespoloną funkcją błędu:

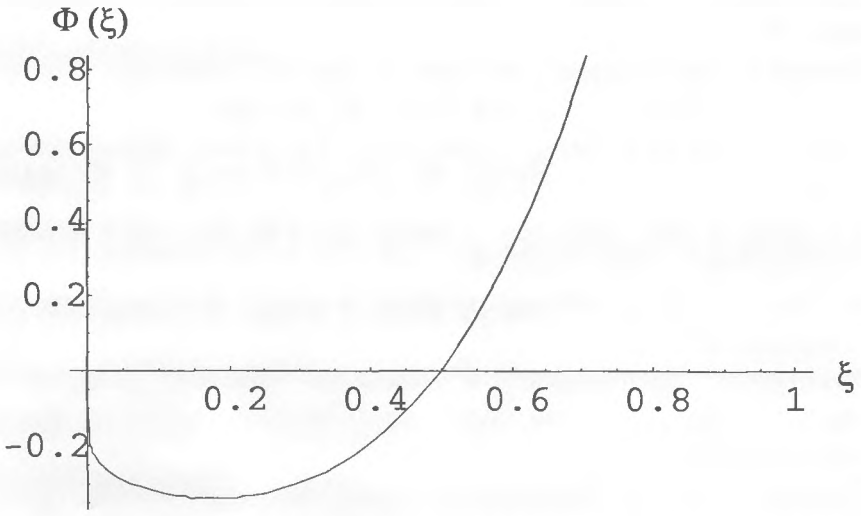
$$erfi(z) = -i\, erf(iz)$$

gdzie erf(z) jest funkcją błędu:

$$erf(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt$$

Mimo bardzo skomplikowanej postaci równania (7), za pomocą programu *Mathematica*® można w analogiczny sposób narysować wykres funkcji Φ (rys. 3)

i wyznaczyć pierwiastek równania (7) $\xi \approx 0.500451$, a następnie $IRR \approx$ ≈ 0.998199.



Rys. 3. Wykres funkcji Φ (wzór (7)) dla potęgowych funkcji przychodów i wydatków

Podsumowanie

Przytoczone przykłady pokazują, że przy wyznaczaniu *IRR* dla inwestycji z gęstymi nośnikami przepływów finansowych mogą się pojawić całki nieelementarne i równania przestępne. Pokazują również, że za pomocą systemów algebry komputerowej (np. *Mathematica*®) można bez większych trudności rozwiązać te równania i wyznaczyć *IRR* przy każdych ustalonych wartościach parametrów, które występują w funkcjach intensywności przychodów i wydatków. Wydaje się, że wspomniane programy mogą się również okazać pomocne w rozwiązaniu problemu, polegającego na sformułowaniu warunków wystarczających na to, by równanie (3) posiadało jeden tylko pierwiastek ξ, i to pierwiastek znajdujący się w przedziale (0, 1).

Literatura

1. Drwal G., Grzymkowski R., Kapusta A., Słota D.: Mathematica 5. Wydawnictwo Pracowni Komputerowej Jacka Skalmierskiego, Gliwice 2004.
2. Jajuga K.: Zarządzanie kapitałem. Wydawnictwo AE im. Oskara Langego we Wrocławiu, Wrocław 1993.

3. Kuratowski K.: Wstęp do teorii mnogości i topologii. Państwowe Wydawnictwo Naukowe, Warszawa 1966.
4. Manikowski A., Tarapata Z.: Ocena projektów gospodarczych. Difin, Warszawa 2002.
5. Piasecki K.: Od arytmetyki handlowej do inżynierii finansowej. Wydawnictwo Akademii Ekonomicznej w Poznaniu, Poznań 2005.
6. Ralston A.: Wstęp do analizy numerycznej. Państwowe Wydawnictwo Naukowe, Warszawa 1971.
7. Tarczyński W., Mojsiewicz M.: Zarządzanie Ryzykiem. Polskie Wydawnictwo Ekonomiczne, Warszawa 2001.
8. Wolfram S.: The *„Mathematica"* Book (5th ed.). Wolfram Media, Inc., Champaign, IL., 2003.
9. Zawadzki H.: *„Mathematica®"* w matematyce finansowej. Obliczanie wewnętrznej stopy zwrotu inwestycji. Zeszyty Naukowe AE, nr 31, Katowice 2004, s.121-133.
10. Zawadzki H. (2004): Matematyczne aspekty obliczania wewnętrznej stopy zwrotu. Zeszyty Naukowe Uniwersytetu Szczecińskiego, nr 389 (Finanse-Rynki finansowe-Ubezpieczenia, nr 2), Szczecin 2004, s. 257-266.

COMPUTATIONAL ASPECTS OF CALCULATING THE INTERNAL RATE OF RETURN FOR INVESTMENTS WITH DENSE SUPPORT OF CASH FLOWS

Summary

In order to find an internal rate of return of investment (IRR) in case of the dense supports of cash flows, we have to solve for ξ an equation:

$$\int_0^T cif_X(t)\xi^t dt = -\int_0^T cof_X(t)\xi^t dt$$

where $T > 0$ denotes the time of living of the project (investment) and functions $cif_X : [0, T] \to \to R^+$ and $cof_X : [0, T] \to R^-$ denote intensities of cash inflows and outflows respectively. It turns out that even in simple cases, integrals that occur in this equation are non elementary integrals and the equation alone is a transcendental equation. In this article we look more closely at two of such cases when intensities of cash flows are exponential or power functions.

Katarzyna Zeug-Żebro

UWAGI O STATYSTYCE BDS I WYKŁADNIKU HURSTA W ODNIESIENIU DO DANYCH GIEŁDOWYCH

Wprowadzenie

W ostatnich czasach coraz częściej pojawiają się nowe metody służące do analizy finansowych szeregów czasowych. Niewątpliwie przyczyną wzrostu zainteresowania tymi metodami jest pogoń za stworzeniem „doskonałej" metody prognozowania cen finansowych, a co za tym idzie – chęć zdobycia ponadprzeciętnego dochodu.

Jedną z grup metod należących do tego nurtu jest teoria chaosu. Ważnym jej dokonaniem jest wykazanie istnienia układów deterministycznych, które w pewnych warunkach zachowują się chaotycznie, nieregularnie. Wynika stąd, że możliwe jest odróżnienie deterministycznych szeregów czasowych od losowych.

Celem artykułu jest próba odróżnienia deterministycznych szeregów czasowych od losowych za pomocą dwóch metod – statystyki BDS i analizy R/S. Dane wykorzystane w opracowaniu pochodzą z GPW w Warszawie z okresu od stycznia 2000 roku. Obliczenia przeprowadzono z użyciem programów napisanych przez autorkę w języku programowania Visual Basic oraz pakietu Microsoft Excel.

1. Statystyka BDS

Statystyka BDS została wprowadzona w 1987 roku przez W. Brocka, W. Decherta i J. Scheinkmana [1]. Opiera się ona na pojęciu całki korelacyjnej i testuje hipotezę, że zbiór danych jest iid (independent, identically distributed,

czyli jest to zbiór niezależnych zmiennych losowych o jednakowym rozkładzie). Jest zatem jedną z metod odróżniania deterministycznych szeregów czasowych od losowych.

Zanim rozwiniemy pojęcie statystyki BDS, określimy najpierw całkę korelacyjną ($C(d, r)$[8]). $C(d, r)$ jest zdefiniowana jako prawdopodobieństwo znalezienia pary wektorów, których odległość od siebie w zrekonstruowanej d-wymiarowej przestrzeni nie jest większa od r:

$$C(d,N,r,t)=\frac{2}{n(n-1)}\sum_{j=1}^{n-1}\sum_{i=j+1}^{n} I\left(r-r_{ij}\right), \quad r>0 \tag{1}$$

gdzie $I(x)$ jest funkcją wskaźnikową (funkcja Heaviside) w postaci:

$$I(a)=\begin{cases}0 & \text{dla } a<0\\ 1 & \text{dla } a\geq 0\end{cases} \tag{2}$$

$n = N - t(d - 1)$ jest liczbą wektorów w d-wymiarowej przestrzeni, N jest liczbą danych, t jest wskaźnikiem opóźnienia, a $r_{ij} = \| x_i - x_j \|$.

Rozwińmy teraz pojęcie testu BDS. Niech F będzie rozkładem (wielowymiarowej) zmiennej X w przestrzeni stanów, a całka korelacyjna ma postać:

$$C(d,r)=\iint I\left(r-\|x-y\|\right)dF(x)dF(y), \quad r>0 \tag{3}$$

Jeśli zmienna X jest iid, to dla:

$$I\left(r-\|x-y\|\right)=\prod_{k=1}^{d} I\left(r-|x_k-y_k|\right)$$

otrzymujemy:

$$C(d,r) = C^d(1,r)$$

gdzie:

$$C(1,r)=\int\left[F(x+r)-F(x-r)\right]dF(x)\equiv C$$

Denker i Keller pokazali, że $C(d,N,r)$ jest estymatorem statystyki U. Natomiast Bock i in., korzystając z teorii statystyki U dla regularnych procesów,

udowodnili, że jeśli $N \to \infty$, to $\sqrt{n}\left[C(d,N,r) - C^d(1,r)\right]$ ma rozkład normalny ze średnią zero i wariancją:

$$\sigma^2(d,r) = 4\left[K^d - C^{2d} + 2\sum_{i=1}^{d-1}\left(K^{d-i}C^{2i} - C^{2d}\right)\right] \tag{4}$$

gdzie:

$$K \equiv \int\left[F(x+r) - F(x-r)\right]^2 dF(x)$$

(zakładamy, że $K > C^2$). Zatem statystyka *BDS* będzie zdefiniowana wzorem:

$$BDS(d,N,r) = \frac{\sqrt{N}}{\sigma(d,r)}\left[C(d,N,r) - C^d(1,r)\right] \tag{5}$$

z rozkładem normalnym. Jeśli rozkład F jest nieznany, to nie możemy wyznaczyć wartości C i K oraz wariancji $\sigma^2(d,r)$. W takim przypadku $C(1,r)$ i $\sigma^2(d,r)$ muszą być oszacowane przez $C(1,N,r,t)$ i:

$$\overline{\sigma}^2 = 4\left\{d(d-1)\overline{C}^{2(d-1)}\left(\overline{K} - \overline{C}^2\right) + \overline{K}^d - \overline{C}^{2d} + 2\sum_{i=1}^{d-1}\left[\overline{C}^{2i}\left(\overline{K}^{d-i} - \overline{C}^{2(d-i)}\right) - d\overline{C}^{2(d-i)}\left(\overline{K} - \overline{C}^2\right)\right]\right\}$$

gdzie $\overline{C} = C(d,N,r,t)$ i:

$$\overline{K} = \frac{6}{n(n-1)(n-2)}\sum_{1\le i\le j\le k\le n} I\left(r - \|y_i - y_j\|\right) I\left(r - \|y_i - y_k\|\right)$$

Zatem statystyka *BDS* przyjmuje postać:

$$BDS(d,N,r) = \frac{\sqrt{n}}{\overline{\sigma}}\left[C(d,N,r,t) - C^d(1,N,r,t)\right] \tag{6}$$

2. Analiza R/S

Wykładnik Hursta [5] jest miarą statystyczną, która pozwala na klasyfikację szeregów czasowych, tj. odróżnienie deterministycznych szeregów czasowych od losowych. Jedną z metod obliczania wykładnika Hursta jest metoda

analizy R/S. Dla szeregu obserwacji $\{x_1,x_2,\ldots,x_N\}$ przebiega ona w następujących etapach:

1. Korzystając z wzoru (7), przekształcamy powyższy szereg w ciąg $m = N - 1$ logarytmicznych stóp zwrotu:

$$y_k = \log\left(\frac{x_{k+1}}{x_k}\right), k = 1, 2, ..., N-1 \tag{7}$$

2. Następnie dzielimy otrzymany ciąg (szereg logarytmicznych stóp zwrotu) na T podciągów złożonych z t elementów (T, t – liczby naturalne, takie że $T \cdot t = m$, gdy m nie ma dzielników, to bierzemy $m' < m$, który ma dzielniki, odrzucając $m - m'$ początkowych wyrazów ciągu). Każdy element podciągu oznaczamy przez m_{ij} dla $i = 1,\ldots,t; j = 1,\ldots,T$. Średnia wartość dla j-tego podciągu wynosi:

$$\bar{y}_j = \frac{\sum_{i=1}^{t} m_{ij}}{t}$$

3. Kolejnym krokiem jest scentrowanie każdego podciągu poprzez odjęcie średniej arytmetycznej:

$$z_{ij} = m_{ij} - \bar{y}_j \tag{8}$$

i zdefiniowanie ciągu sum częściowych z_{ij}:

$$q_{ij} = \sum_{l=1}^{t} z_{lj}, i = 1, 2, ..., t, \text{ j} = 1, 2, ..., T \tag{9}$$

Zauważmy, że q_i jest skumulowanym odchyleniem od wartości średniej dla pierwszych j wartości w przedziale i.

4. Następnie obliczamy rozstępy skumulowanych szeregów czasowych według wzoru:

$$R_j = \max(q_{ij}) - \min(q_{ij}) \tag{10}$$

5. Ostatnim krokiem jest obliczenie dla każdego skumulowanego szeregu czasowego rozstępu przeskalowanego, tzn. podzielenie rozstępu przez odchylenie standardowe tego szeregu:

$$\alpha_{jt} = R_j/S_j \tag{11}$$

gdzie:

$$S_j = \sqrt{\frac{1}{t}\sum_{i=1}^{t} z_{ij}^2}$$

6. Powyższą procedurę przeprowadza się dla różnych długości szeregu czasowego t. W ten sposób otrzymujemy zależność wielkości R/S od długości szeregu t.

Aby wyznaczyć wykładnik Hursta, należy zlogarytmować następującą zależność:

$$E(R/S) = ct^H \tag{12}$$

gdzie H jest wykładnikiem Hursta, c jest stałą, a $E(R/S)$ jest wartością oczekiwaną przeskalowanego zakresu:

$$\ln E(R/S) = \ln c + H \ln t \tag{13}$$

Wykładnik Hursta jest współczynnikiem kierunkowym regresji liniowej.

Ponieważ wykładnik Hursta służy do odróżniania deterministycznych szeregów czasowych od losowych, zostaną omówione teraz zasady klasyfikacji. Szeregi czasowe można podzielić na trzy klasy w zależności od ich wartości. Jeśli szereg jest szeregiem losowym, to wartość wykładnika Hursta jest równa 0,5, czyli w takim szeregu elementy występujące po sobie są niezależne i wartość teraźniejsza nie ma wpływu na wartość przyszłą.

Wartość wykładnika Hursta z przedziału (0,5;1) wskazuje nam szeregi persystentne, czyli takie szeregi losowe, w których obserwujemy trend (ciągle wzmacniający się). Szeregi takie zachowują pamięć nie tylko krótkookresową, ale i sięgającą daleko wstecz, jednak zbyt odległe elementy mają niewielki wpływ na teraźniejszość.

Do ostatniej klasy należą wartości wykładnika Hursta z przedziału (0;0,5). Wartości takie wskazują nam szeregi antypersystentne. Współczynnik korelacji ma w takim przypadku wartość ujemną. Jeśli wykładnik H jest bliski zeru, to współczynnik korelacji zmniejsza swoją wartość i obserwuje się w szeregu wzrost szumu.

3. Wyniki analizy empirycznej

Przeprowadzone badania pozwoliły za pomocą statystyki BDS i analizy R/S, zweryfikować hipotezę o deterministycznym charakterze badanego szeregu czasowego.

Pod uwagę wzięto indeks WIG i siedem spółek, których akcje były notowane na Giełdzie Papierów Wartościowych w Warszawie co najmniej od stycznia 2000 roku. Razem przeanalizowano 1480 obserwacji, którymi były dzienne stopy zwrotu:

$$R_t = \ln P_t - \ln P_{t-1}$$

Obliczając statystykę BDS, przyjmowano różne wymiary zanurzenia (d = = 2, 3, 4, 5) oraz różne wartości r (0.25σ;0.5σ;0.75σ)[5], gdzie σ jest odchyleniem standardowym z próby. Otrzymane wyniki przedstawia tabela 1.

Tabela 1

Zestawienie wartości statystyki BDS dla szeregów czasowych utworzonych z notowań wybranych spółek

Spółka	d	r			
		$0{,}25\sigma$	$0{,}5\sigma$	$0{,}75\sigma$	σ
1	2	3	4	5	6
BPHPBK	2	4,1223	3,2331	2,7364	2,4522
	3	6,3243	5,7865	5,5564	4,2343
	4	7,9876	5,6453	4,02343	3,8976
	5	9,0032	8,9987	8,0324	7,6754
BZWBK	2	2,3426	2,4532	1,8976	1,0987
	3	3,2673	3,0324	2,7654	2,4352
	4	5,8763	5,0982	4,6547	4,0478
	5	7,8395	7,0326	7,1234	6,9328
DĘBICA	2	2,3454	1,9876	1,4673	1,8761
	3	2,7865	2,6753	2,8975	1,9871
	4	3,0123	2,9876	2,5674	1,8765
	5	3,9876	3,7865	3,9876	3,0126
INGBSK	2	5,2963	4,4352	4,0987	3,7654
	3	6,7895	6,1278	5,7589	5,1098
	4	7,3426	7.0023	6,6784	6,0987
	5	8,6543	7,9876	7,1987	6,9804
JUTRZENKA	2	4,3784	3,47865	3,8765	3,0002
	3	6,8976	5,7856	3,8987	2,7466
	4	9,6488	9,0987	6,7466	4,7489
	5	11,9773	10,3643	9,8889	8,4648
OPTIMUS	2	3,4785	2,4751	1,9478	1,0674
	3	5,7321	4,8335	4,2382	3,1231
	4	7,7384	6,7368	6,7822	5,8722
	5	8,8432	7,8732	6,7389	6,1231

cd. tabeli 1

1	2	3	4	5	6
RAFAKO	2	7,3454	6,3729	8,7249	5,8232
	3	9,2398	8,3793	7,8723	6,9121
	4	11,8923	9,9823	9,0281	8,9841
	5	13,7841	11,8273	10,7823	9,8191
WIG	2	7,2419	6,1883	5,9821	5,0184
	3	8,7893	6,9148	6,1913	4,8778
	4	10,8231	8,1431	7,1831	5,8392
	5	12,7821	9,3144	8,9713	8,4189

Analizując otrzymane wyniki, odrzucamy hipotezę, że badane szeregi są ciągami wartości niezależnych zmiennych losowych o jednakowym rozkładzie, czyli badane szeregi czasowe nie są czysto losowe. Jednakże odrzucenie przez test BDS tej hipotezy nie jest równoznaczne z założeniem, że badany szereg jest deterministyczny. Informację tę należy przyjąć jako wstęp do dalszych badań.

W drugiej części badań empirycznych zastosowano analizę R/S. Tak jak w poprzednim przypadku dysponowano 1480 obserwacjami. Pod uwagę wzięto różne długości podszeregów czasowych. Po obliczeniu wartości R/S ustalono wartość wykładnika Hursta poprzez wyznaczenie współczynnika kierunkowego prostej regresji (ln *R/S* względem ln *t*). Wyniki obliczeń przedstawia tabela 2.

Tabela 2

Wartość wykładnika Hursta dla szeregów czasowych utworzonych z notowań wybranych spółek

SPÓŁKA	WYKŁADNIK HURSTA
BPHPBK	0,4491
BZWBK	0,511
DĘBICA	0,5876
INGBSK	0,5163
JUTRZENKA	0,5381
OPTIMUS	0,4806
RAFAKO	0,5271
WIG	0,4958

Na podstawie otrzymanych wyników można wnioskować, że szeregi czasowe utworzone z cen akcji BZWBK, Dębica, INGBSK, Jutrzenka i Rafako charakteryzują się pamięcią krótkookresową (wartość wykładnika Hursta jest większa od 0,5). Zatem w ich szeregach czasowych utworzonych ze stóp zwrotu jest możliwe istnienie pewnej deterministycznej struktury.

Podsumowanie

W opracowaniu podjęto próbę odróżnienia deterministycznych szeregów czasowych od losowych, korzystając z dwóch metod: statystyki BDS i analizy R/S. Badania zostały przeprowadzone dla indeksu WIG i siedmiu spółek (BPHPBK, BZWBK, Dębica, INGBSK, Jutrzenka, Optimus, Rafako) notowanych na GPW w Warszawie. Rozważane szeregi składały się z cen zamknięcia i pochodziły z okresu od stycznia 2000 roku. Badania przeprowadzono korzystając z programu napisanego przez autorkę w języku programowania Visual Basic oraz z pakietu Excel.

Na podstawie przeprowadzonych badań można zauważyć, że bardzo trudno jest jednoznacznie określić, czy szeregi finansowe są deterministyczne, czy losowe. Możliwość istnienia pewnej chaotycznej, deterministycznej struktury należy traktować z ostrożnością, gdyż długość rozpatrywanych szeregów obserwacji może być zbyt krótka. Informację taką należy potraktować jako zachętę do dalszych badań z wykorzystaniem innych metod odróżniania deterministycznych szeregów czasowych od losowych. Być może potwierdzą one lub też zaprzeczą istnieniu chaosu na giełdzie.

Literatura

1. Brock W.A., Dechert W.D., Scheinkman J.A.: A Test for Independence Based on Correlation Dimension. SSRI Working Paper No. 8702, Department of Economics, University of Winsconsin, Madison, WI. 1987.
2. Grassberger P., Procaccia I.: Characterization of Strange Attractors. „Phys. Rev. Lett." 1983, Vol. 50, pp. 346-349.
3. Grassberger P., Procaccia I.: Measuring the Strangeness of Strange Attractors. „Physica D" 1983a.
4. Kim H.S, Eykholt R., Salas J.D.: Nonlinear Dynamics, Delay Time, and Embedding Windows. „Physica D 127" 1999.
5. Peters E.E.: Chaos and Order in the Capital Markets. Wiley, New York 1991.

6. Scheinkman J.A., LeBaron B.: Nonlinear Dynamics and Stock Returns. „The Journal of Business" 1989a, Vol. 62, No. 3, pp. 311-337.
7. Schuster H.G.: Chaos deterministyczny. PWN, Warszawa 1995.
8. Zawadzki H.: Chaotyczne systemy dynamiczne. AE, Katowice 1996.

BDS STATISTIC AND R/S ANALYSIS – THE METHOD DISTINGUISHING RANDOM AND DETERMINISTIC SYSTEMS

Summary

In this paper we discuss some recent techniques used in distinguishing between probabilistic and deterministic behavior in stock price: the BDS statistic and R/S analysis. Our data set is composed of daily data obtained from GPW in Warsaw for index WIG and seven company.

Informacja o Katedrze Matematyki

Katedra Matematyki został[a] ra Akademii Ekonomicznej w Katow[icach] ez wyodrębnienie z Katedry Ekonom[etrii] ze organizacyjnej Wydziału Zarząd[zania] zy zakłady, a mianowicie: Zakład [Statystyki] ad Ekonomii Matematycznej i Zakład Matematyki.

W Katedrze jest zatrudnionych łącznie piętnaście osób, w tym czterech samodzielnych pracowników naukowych mających stopień doktora habilitowanego, jedna osoba z tytułem naukowym profesora nauk ekonomicznych, a dwie są zatrudnione na stanowisku profesora Akademii Ekonomicznej w Katowicach.

Zainteresowania naukowe pracowników Katedry Matematyki obejmują problemy związane z zastosowaniami metod ilościowych w naukach ekonomicznych, a w szczególności: zastosowania analizy matematycznej i algebry liniowej w zarządzaniu i ekonomii, ekonomię matematyczną, metody numeryczne, równania różnicowe, stochastyczne równania różniczkowe, rachunek prawdopodobieństwa, statystykę społeczną, statystykę i modelowanie przestrzenne, pola losowe, gry statystyczne, wielowymiarową analizę porównawczą, metody taksononomiczne, niezawodność w systemach ekonomicznych, teorię masowej obsługi, chaotyczne systemy dynamiczne, analizy fraktalne, teorię użyteczności, teorię podejmowania decyzji, makromodele optymalizacyjne, badania operacyjne, programowanie matematyczne, programowanie liniowe, programowanie dynamiczne, programowanie stochastyczne, teorię grafów i sieci, teorię gier, programowanie wielokryterialne, programowanie dyskretne, teorię rozwiązań uogólnionych w programowaniu matematycznym, teorię korekcji zadań optymalizacyjnych, agregację modeli optymalizacyjnych, algorytmy genetyczne, sieci neuronowe, metody badań rozwoju gospodarczego, modelowanie procesów ekonomicznych, modelowanie zjawisk demograficznych, integrację europejską, matematykę finansową, inżynierię finansową, analizę ekonomiczną, analizę finansową przedsiębiorstw, wizualizację w nauczaniu matematyki i przedmiotów pokrewnych, programowanie komputerów.

ISBN 978-83-7246-544-3