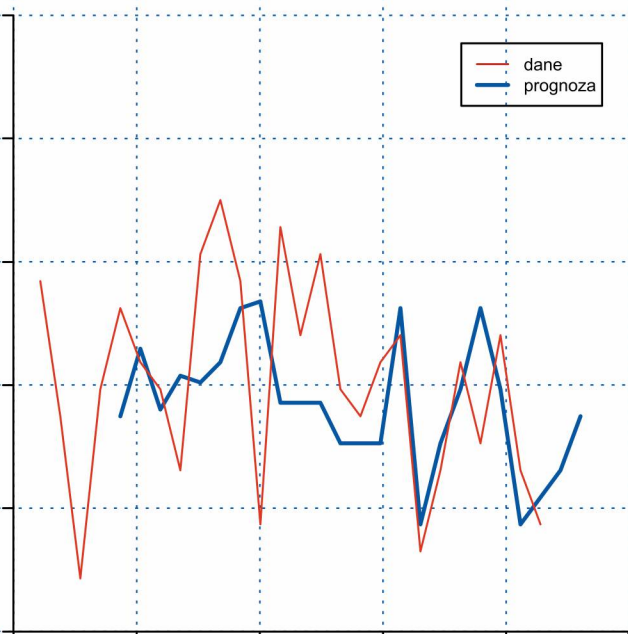


Adam Kucharski

# Prognozowanie szeregów czasowych metodami ewolucyjnymi



Adam Kucharski

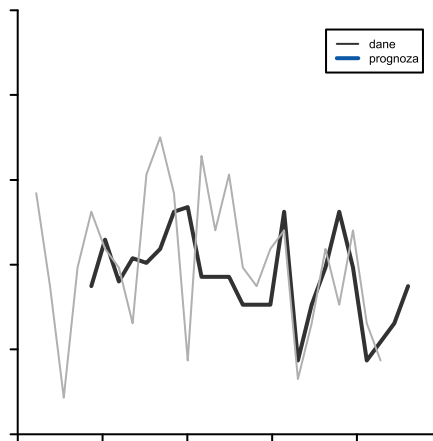
**Prognozowanie**  
szeregów czasowych  
metodami ewolucyjnymi



WYDAWNICTWA  
UNIwersytetu  
ŁÓDZKIEGO

Adam Kucharski

# Prognozowanie szeregów czasowych metodami ewolucyjnymi



WYDAWNICTWO  
UNIwersYTETU  
ŁÓDZKIEGO

ŁÓDŹ 2013

Adam Kucharski – Katedra Badań Operacyjnych, Wydział Ekonomiczno-Socjologiczny  
Uniwersytet Łódzki, 90-214 Łódź, ul. Rewolucji 1905 r. nr 41/43

RECENZENT

*Tadeusz Trzaskalik*

SKŁAD I ŁAMANIE

*Adam Kucharski*

PROJEKT OKŁADKI

*Łukasz Orzechowski*

© Copyright by Uniwersytet Łódzki, Łódź 2013

Wydane przez Wydawnictwo Uniwersytetu Łódzkiego  
Wydanie I. W.06283.13.0.H

ISBN (wersja drukowana) 978-83-7525-936-0  
ISBN (ebook) 978-83-7969-163-0

Wydawnictwo Uniwersytetu Łódzkiego  
90-131 Łódź, ul. Lindleya 8  
www.wydawnictwo.uni.lodz.pl  
e-mail: ksiegarnia@uni.lodz.pl  
tel. (42) 665 58 63, faks (42) 665 58 62

# Spis treści

Wstęp . . . . .	7
Rozdział 1. Algorytmy genetyczne – charakterystyka metody . . . . .	11
1.1. Klasyczny algorytm genetyczny . . . . .	11
1.1.1. Matematyczne podstawy algorytmów genetycznych . . . . .	16
1.1.2. Przykład działania klasycznego algorytmu genetycznego . . . . .	21
1.2. Przegląd modyfikacji wpływających na konstrukcję chromosomu . . . . .	30
1.2.1. Zmiany reprezentacji genotypu . . . . .	30
1.2.2. Działania wpływające na porządek genów w chromosomie . . . . .	33
1.2.3. Kodowanie rzeczywiste chromosomu . . . . .	36
1.3. Przegląd modyfikacji operatorów algorytmu . . . . .	40
1.3.1. Nieklasyczne metody selekcji chromosomów . . . . .	40
1.3.2. Model uogólnionego krzyżowania . . . . .	47
1.3.3. Wykorzystanie zjawiska nisz i gatunków . . . . .	48
1.4. Przegląd modyfikacji zmieniających postać funkcji przystosowania . . . . .	53
1.4.1. Skalowanie funkcji przystosowania . . . . .	53
1.4.2. Algorytmy ze zmienną liczebnością populacji . . . . .	55
1.4.3. Zadania z ograniczeniami . . . . .	57
1.4.4. Zastosowanie algorytmów genetycznych do optymalizacji zadań wielokryteriowych . . . . .	59
1.5. Strategie ewolucyjne . . . . .	61
1.6. Algorytm genetyczny jako element algorytmów ewolucyjnych . . . . .	63
Rozdział 2. Prognozowanie – własności i metody . . . . .	65
2.1. Podstawowe pojęcia . . . . .	65
2.2. Dekompozycja szeregu czasowego . . . . .	71
2.3. Prognozowanie przy pomocy metod mechanicznych . . . . .	75
2.3.1. Metody naiwne . . . . .	75
2.3.2. Metoda średniej ruchomej . . . . .	77
2.3.3. Metoda wygładzania wykładniczego Browna . . . . .	81
2.3.4. Metoda Holta . . . . .	82
2.3.5. Metoda Wintersa . . . . .	84
2.4. Błędy prognoz <i>ex post</i> . . . . .	88
2.5. Charakterystyka metod analitycznych . . . . .	95
Rozdział 3. Algorytm genetyczny w prognozowaniu . . . . .	99
3.1. Progностyczny algorytm genetyczny . . . . .	99

---

3.1.1. Uwzględnianie trendu w PAG . . . . .	109
3.1.2. Uwzględnianie sezonowości . . . . .	111
3.2. Algorytm bazujący na średniej ruchomej . . . . .	115
3.3. Prognozowanie kaskadowe . . . . .	119
3.3.1. Podejście bayesowskie w prognozowaniu kaskadowym . . . . .	123
3.3.2. Prognozowanie kaskadowe – przykład . . . . .	127
Rozdział 4. Prognozy wybranych kategorii przy pomocy PAG . . . . .	133
4.1. Prognozy <i>ex post</i> i <i>ex ante</i> wybranych szeregów danych . . . . .	133
4.1.1. Prognozy notowań i wolumenu obrotów danych giełdowych . . . . .	133
4.1.2. Prognozy notowań kursów walut . . . . .	156
4.1.3. Prognozy IRDN oraz inflacji . . . . .	169
4.1.4. PAG a zmienne losowe . . . . .	177
4.2. Efektywność PAG podczas przeszukiwania zbioru wariantów prognostycznych . . . . .	181
4.2.1. Zmiany przystosowania najlepszego chromosomu populacji . . . . .	183
4.2.2. Zmiany średniego przystosowania chromosomów populacji . . . . .	197
4.3. Dobór parametrów PAG . . . . .	208
Zakończenie . . . . .	215
Bibliografia . . . . .	217

## Wstęp

Z prognozami stykamy się częściej niż może to się wydawać na pierwszy rzut oka. Pierwsze skojarzenie zwykle wiąże się z przepowiadaniem pogody, a tymczasem z prognoz (niekiedy mówimy o przewidywaniach) korzystamy na co dzień i to często nieświadomie. Wyobraźmy sobie, że zastanawiamy się nad wyborem trasy pokonywanej samochodem w mieście. Nim wyruszymy w drogę rozważamy na przykład czy i gdzie możemy trafić na korki o określonej porze dnia. Uwzględnienie spodziewanych utrudnień w ruchu spowodowanych godzinami szczytu stanowi właśnie przykład nieświadomego przewidywania. Nie planujemy jednak zajmować się bezrefleksyjnymi zachowaniami. Planujemy sięgnąć po metody naukowe, oparte na świadomym, racjonalnym działaniu. Spośród różnych obszarów zainteresowania prognozowania jako nauki zdecydowaliśmy się na predykcję na podstawie szeregów czasowych. Skłoniło nas do tego kilka przyczyn.

Po pierwsze, w przypadku niektórych zjawisk decyzje należy podejmować na podstawie samych szeregów czasowych gdyż albo nie ma dość czasu na poszukiwanie zewnętrznych czynników wpływających na dane zdarzenie, albo takowych nie da się wskazać w jednoznaczny sposób. Na przykład decydent zainteresowany podjęciem szybkiej decyzji dotyczącej rynku finansowego może nie dysponować wystarczającą ilością czasu aby zbudować model ekonometryczny. Ma dostęp do szeregów notowań i niejednokrotnie tyle musi mu wystarczyć.

Po drugie uważamy, że wciąż nie powiedziano ostatniego słowa jeśli chodzi o metody, które można w przypadku szeregów czasowych wykorzystać. Mamy tu na myśli zarówno modyfikacje istniejących jak i opracowywanie nowych metod. Istnieje już wprawdzie całkiem całkiem pokazny ich zbiór, ale wciąż można znaleźć słabo zbadane obszary zastosowań. Wiążą się one m.in. z uchylaniem założeń ograniczających zastosowania metod do szeregów czasowych o ściśle określonej dekompozycji. Dowiedziemy, że możliwym jest uogólnienie restrykcyjnych warunków stosowalności wybranych metod, aby osiągnąć większą elastyczność ich wykorzystania.

Po trzecie, opieranie się w prognozach na informacjach na temat mechanizmów odpowiedzialnych za zachowanie zjawiska, które obecne są w szeregu czasowym bynajmniej nie wyklucza tej drogi postępowania w porównaniu



z narzędziami o bardziej skomplikowanej konstrukcji. W konfrontacji z nimi (na przykład jedno- i wielorównaniowymi modelami ekonometrycznymi) takie postępowanie wydaje się zbyt uproszczone, żeby nie powiedzieć „trywialne”. Praktyka jednak pokazuje, że metoda prostsza nie oznacza automatycznie metody gorszej. Dobrze to ilustruje porównanie metod naiwnych i wygładzania wykładniczego. W przypadku tych drugich optymalizacja parametrów potrafi doprowadzić do bardzo dokładnego dopasowania prognoz do danych z próby co jednak wcale nie gwarantuje równie wysokiej trafności predykcji poza samą próbą. Metody naiwne potrafią prognozować te same dane ze zbliżoną dokładnością a są o wiele łatwiejsze do zastosowania.

Kolejna przyczyna wiąże się z planowanym wykorzystaniem algorytmów genetycznych i ewolucyjnych. Wywodzą się one ze starej jak sama ludzkość chęci naśladowania lub odwzorowywania procesów, zjawisk i zdarzeń obserwowanych w naturze. Należy zauważyć, że tego typu narzędzia, choć rozwijane od kilku dziesięcioleci, wciąż słabo eksplorują niektóre obszary potencjalnych zastosowań. Zaliczymy do nich m.in. prognozowanie. Przeanalizowana na potrzeby tej pracy literatura dowodzi, że naprawdę niewiele publikacji traktuje o połączeniu algorytmów genetycznych i ewolucyjnych z prognozowaniem. Jeśli już do tego dochodzi to często celem staje się tak naprawdę opracowanie generatora reguł decyzyjnych operującego na już istniejących metodach. Przesadą byłoby oczywiście stwierdzenie, że wkraczamy w tej książce na zupełnie dziewicze tereny. Tym niemniej warto podkreślić, że zaproponowane w niej wykorzystanie algorytmu genetycznego do zastosowań prognostycznych stanowi oryginalny wkład Autora. Trzymana w rękę praca stanowi próbę przetrzepania pomostu pomiędzy prognozowaniem niestrukturalnym a wybranymi metodami optymalizacji heurystycznej. Postawiliśmy sobie za cel udowodnienie, że takie połączenie jest nie tylko możliwe, ale że niesie ze sobą również korzyści.

Nie mamy zamiaru deprecjonować stworzonych do tej pory istniejących metod i narzędzi. W istocie stanowią one cenną inspirację. Zdajemy sobie jednak sprawę z ich niedoskonałości w określonych okolicznościach. Z punktu widzenia prognozowania interesować nas będą szczególnie takie sytuacje, w których metody (umownie nazwane klasycznymi) mają kłopoty z trafnością predykcji i odwzorowaniem mechanizmów stojących za zachowaniem szeregu.

Podstawowym celem jaki sobie stawiamy jest opracowanie zmodyfikowanego algorytmu genetycznego, który miałby posłużyć do prognozowania szeregów czasowych charakteryzujących się skomplikowaną i/lub w jakiś sposób nietypową dekompozycją. Istnieje wiele metod przystosowanych do szeregów o konkretnym zachowaniu, ale istnieje też wiele zachowań nieprzystających do założeń tych metod. Uważamy, że elastyczność będąca

charakterystyczną cechą metod optymalizacji heurystycznej pozwala przezwyciężyć te ograniczenia.

Od strony algorytmu genetycznego lub ewolucyjnego chcieliśmy przedstawić, w jaki sposób można uwzględnić w nim wiedzę specyficzną dla zadania oraz jak poprawia to działanie systemu ewolucyjnego (nazwę cytujemy za Z. Michalewiczem (Michalewicz 2003)). W tym celu opracowaliśmy algorytm, który nazwaliśmy *Prognostycznym Algorytmem Genetycznym* (PAG) będący hybrydą algorytmów genetycznych i ewolucyjnych przystosowaną do tworzenia prognoz na podstawie szeregów czasowych.

Ponadto postanowiliśmy poszukać rozwiązania innej bolączki występującej w przypadku prognozowania niestrukturalnego. Metody zaliczane do tej grupy nie uwzględniają długookresowych procesów obecnych w szeregach czasowych, zwykle zadowalając się wykorzystywaniem obserwacji pochodzących z niedalekiej przeszłości. Problem w ich wypadku stanowią również prognozy *ex ante*, które sięgają bardzo blisko w przyszłość. Dlatego celem dodatkowym uczyniliśmy znalezienie rozwiązania tych wad właśnie przy pomocy PAG.

Pracę podzieliliśmy na cztery rozdziały. Pierwszy z nich poświęcony został zaprezentowaniu idei algorytmów genetycznych wraz ze stojącymi za nimi definicjami i twierdzeniami matematycznymi. Ponadto zawarliśmy tam opis wybranych modyfikacji, którym poddano te algorytmy w ciągu lat badań i analiz. Zmiany objęły takie elementy jak sposób kodowania, metody selekcji i inne. Konsekwencją wprowadzanych modyfikacji stało się uogólnienie algorytmów genetycznych do algorytmów ewolucyjnych. Na zakończenie tego rozdziału podjęliśmy próbę wskazania występujących między nimi różnic i podobieństw.

Rozdział drugi traktuje o prognozowaniu. Przybliżyliśmy w nim założenia, funkcje i cechy charakterystyczne tej dziedziny nauki. W dalszej części skupiamy się na prezentacji metod prognostycznych przydatnych podczas predykcji opartej na szeregach czasowych. Zdecydowaliśmy się wyjść poza spotykane zwykle w literaturze samoistne ograniczenie do narzędzi o podręcznikowym charakterze. Oprócz nich zawarliśmy w tej części pracy przegląd wyników najnowszych badań dotyczących różnych punktów widzenia na metodologię prognoz oraz sposoby oceniania ich dokładności.

Trzeci rozdział poświęciliśmy połączeniu algorytmu genetycznego i prognozowania w jeden prognostyczny algorytm genetyczny. Omówiliśmy jego konstrukcję wraz z przykładami ilustrującymi zachowanie dla różnych typów szeregów czasowych. W polskiej literaturze brak prac o podobnej tematyce, a i w zagranicznej nie ma ich zbyt wiele. Zawarliśmy jednak omówienie prac anglojęzycznych, w których autorzy zmierzali się ze wspomnianym zagadnieniem w większym lub mniejszym zakresie. Należy w tym miejscu podkreślić, że PAG stanowi oryginalny wkład badawczy Autora pracy, będący wynikiem

serii przemysłów i badań. Pozostali autorzy skupiali się na jednym z elementów – algorytmie genetycznym lub ewolucyjnym, czy też jednym z aspektów prognozowania. W porównaniu z ich opracowaniami przedstawiamy metodę o bardziej kompleksowym charakterze.

W rozdziale czwartym zdecydowaliśmy się zamieścić wyniki obliczeń uzyskanych przy pomocy prognostycznego algorytmu genetycznego tworzącego tzw. prognozy kaskadowe dla szeregów czasowych pochodzących z różnych źródeł i prezentujących wybrane obszary działalności gospodarczej. Szeregi te ponadto charakteryzują się odmienną dekompozycją, utrudniającą lub wręcz uniemożliwiającą wykorzystanie klasycznych metod. PAG posłużył w ich wypadku do wyznaczenia prognoz *ex post* i *ex ante*. Ostatnie z wymienionych oceniliśmy również jako prognozy wygasłe, aby wykazać użyteczność przyjętego podejścia. Poza tym rozdział zawiera analizę efektywności funkcjonowania opisanego algorytmu pod kątem efektywności poszukiwania rozwiązania.

Różne sposoby uwzględniania specyfiki właściwej dla ewolucji występującej w przyrodzie upodobniły się do siebie w miarę upływu lat. Strategie ewolucyjne, algorytmy genetyczne czy programowanie ewolucyjne przejmowały nawzajem od siebie charakterystyczne dla nich rozwiązania. Czyniło je to coraz bardziej podobnymi do siebie zaś ich czyste, zbliżone do początkowych, formy są dziś rzadko spotykane. Obecnie każdą z wymienionych technik optymalizacji heurystycznej (wraz z kilkoma innymi) łączy się zazwyczaj w postaci jednego nurtu – algorytmów ewolucyjnych. Sam algorytm genetyczny uważany jest za część szerszego zbioru algorytmów ewolucyjnych (szerzej na ten temat piszemy w rozdziale pierwszym).

Z tego powodu proponowany algorytm należy rozpatrywać w szerszym kontekście. Jako wyróżnik klasycznego algorytmu genetycznego opracowanego przez J. Hollanda przyjmuje się m.in. binarne kodowanie i selekcję metodą koła ruletki. Problem polega na określeniu granicy, po przekroczeniu której algorytm genetyczny staje się algorytmem ewolucyjnym. W naszej procedurze, którą stworzyliśmy do celów prognostycznych można wskazać związki z propozycjami Hollanda. Na przykład użyty operator krzyżowania ma konstrukcję analogiczną do zastosowanej przez tego autora. Z drugiej strony chromosomy posiadają kodowanie rzeczywiste co łączy je z algorytmem ewolucyjnym. Trudno zatem jednoznacznie zaliczyć przyjętą przez nas drogę postępowania podczas predykcji szeregów czasowych do jednej, konkretnej rodziny metod. Zdecydowaliśmy więc, aby dla ujednolicenia terminologii określać algorytm właśnie mianem genetycznego.