

Anna Romowicz

Akademia Ekonomiczna w Katowicach

O MOŻLIWOŚCI ZASTOSOWANIA ALGORYTMU GENETYCZNEGO W PROGNOZOWANIU

1. Wstęp

W prezentowanym artykule zostanie przedstawiona propozycja szacowania parametrów α , β , γ trójparametrowego modelu Holta przy wykorzystaniu algorytmu genetycznego. Model liniowy Holta stosowany jest do wygładzania szeregu czasowego, w którym występują i tendencja rozwojowa, i wahania przypadkowe. Model ten jest bardziej elastyczny niż prosty model wygładzania wykładniczego ze względu na występowanie większej liczby parametrów, jednakże problemem związanym ze stosowaniem tego modelu jest określenie wartości parametrów wygładzania α , β i γ . Ich poszukiwania obejmują najczęściej wykonanie serii eksperymentów komputerowych, polegających na stosowaniu różnych kombinacji wartości tych parametrów, a następnie na wyborze tej, która minimalizuje błąd prognoz wygaśłych.

Jedną z proponowanych metod doboru parametrów wygładzania dla trójparametrowego modelu Holta jest wprowadzenie adaptacyjnej procedury dostrajania parametrów wygładzania. Procedura ta zostanie przedstawiona w dalszej części artykułu. Daje ona dobre rezultaty przy odpowiednio długich szeregach czasowych, jednakże może być niesatysfakcjonująca przy niedostatecznie dużej liczbie obserwacji. Ważny wtedy staje się odpowiedni dobór początkowych wartości parametrów α , β i γ . Propozycją przedstawioną w artykule jest ustalenie tychże wartości poprzez wykorzystanie algorytmu genetycznego, co zostanie opisane dokładniej w dalszej części. Na zakończenie zostaną przedstawione wyniki otrzymane po zastosowaniu procedury dla danych dotyczących cen miedzi na Londyńskiej Giełdzie Metali (LME).

2. Model liniowy Holta z trzema parametrami

Model składa się z następujących równań:

$$S_t = \alpha X_t + (1 - \alpha)(S_{t-1} + T_{t-1} + R_{t-1}),$$

$$T_t = \beta dS_t + (1 - \beta)T_{t-1},$$

$$R_t = \gamma d^2 S_t + (1 - \gamma)R_{t-1},$$

gdzie: X_t – szereg czasowy, na podstawie którego wyznaczana jest prognoza,

S_t – szereg wygładzony,

dS_t – różnica pierwszego rzędu, czyli $dS_t = S_t - S_{t-1}$,

$d^2 S_t$ – różnica drugiego rzędu, czyli $d^2 S_t = dS_t - dS_{t-1}$,

T_t – ocena przyrostu trendu liniowego,

R_t – ocena przyrostu trendu kwadratowego,

α, β, γ – parametry modelu o wartościach z przedziału $[0, 1]$.

Prognoza wartości $x_t^*(m)$ z wyprzedzeniem m -krokowym wyraża się wzorem:

$$x_t^*(m) = S_t + mT_t + \frac{1}{2} m^2 R_t.$$

Parametry α, β, γ powinny być tak dobrane, aby minimalizowały średni błąd prognoz wygasłych. Podczas doboru tych parametrów trzeba zwrócić uwagę na dwa kryteria. Po pierwsze, model musi nadążać za zmianami wartości w czasie, a po drugie, powinien odróżniać wahania przypadkowe szeregu X_t od zmian wartości prognoz x_t^* .

3. Adaptacyjna procedura dostrajania parametrów wygładzania

Jedną z metod wyznaczania parametrów α, β, γ jest rozwinięcie trójparametrowego modelu Holta poprzez wprowadzenie procedury dostrajania parametrów wygładzania. Polega ona na wykonaniu kolejnych kroków:

- a) Dla każdej chwili t oblicza się zestaw prognoz (z wyprzedzeniem n) dla trzech wartości każdego parametru:

$$\alpha = \{\alpha_0 - \Delta\alpha^-; \alpha_0; \alpha_0 + \Delta\alpha^+\},$$

$$\beta = \{\beta_0 - \Delta\beta^-; \beta_0; \beta_0 + \Delta\beta^+\},$$

$$\gamma = \{\gamma_0 - \Delta\gamma^-; \gamma_0; \gamma_0 + \Delta\gamma^+\},$$

takich, że

$$\Delta\alpha^+ = \begin{cases} \varepsilon_\alpha(1-\alpha) & \text{dla } \alpha \geq 0,5 \\ \varepsilon_\alpha(1-\alpha)/4 & \text{dla } \alpha < 0,5 \end{cases}$$

oraz

$$\Delta\alpha^- = \begin{cases} \varepsilon_\alpha\alpha/2 & \text{dla } \alpha \geq 0,5 \\ \varepsilon_\alpha\alpha & \text{dla } \alpha < 0,5 \end{cases}$$

gdzie $\varepsilon_\alpha \in (0, 1)$.

Wartości $\Delta\beta$ oraz $\Delta\gamma$ wyznacza się analogicznie do $\Delta\alpha$.

- b) W każdym momencie t wybierany jest zestaw parametrów $\{\alpha', \beta', \gamma'\}$, dla których absolutny błąd prognoz wygasłych jest najmniejszy dla wyprzedzenia o n okresów.
c) Jeżeli błąd najlepszej prognozy spełnia relację:

$$\left| \frac{e(\alpha', \beta', \gamma')}{e(\alpha_0, \beta_0, \gamma_0)} - 1 \right| < \lambda,$$

gdzie $\lambda > 0$ jest zadany progmem, to do dalszych obliczeń wprowadza się nowy zestaw parametrów:

$$\alpha_0^\# = \nu \cdot \alpha_0 + (1-\nu)\alpha',$$

$$\beta_0^\# = \nu \cdot \beta_0 + (1-\nu)\beta',$$

$$\gamma_0^\# = \nu \cdot \gamma_0 + (1-\nu)\gamma'.$$

Jeżeli błąd najlepszej prognozy nie spełnia podanej relacji, to parametry pozostają niezmienione, tzn. wynoszą $\{\alpha_0, \beta_0, \gamma_0\}$.

Tabela 1. Przykładowe obliczenia programu realizującego metodę dostrajania parametrów

$\alpha =$	0,80	$\beta =$	0,30	$\gamma =$	0,50		$m =$	1,00		
$\varepsilon_\alpha =$	0,25	$\varepsilon_\beta =$	0,25	$\varepsilon_\gamma =$	0,25	$\lambda =$	0,40			

t	x_t	s_t	t_t	r_t	$\hat{x}_t^*$	e_t	$\hat{x}_{td}$	α	β	γ
1	3133,50	3133,50	-61,50					0,80	0,30	0,50
2	3072,00	3072,00	-61,50	84,50	3052,75			0,80	0,30	0,50
3	3095,00	3095,00	-32,45	84,50	3104,80	561,84	3104,80	0,80	0,34	0,50
4	3128,50	3132,21	-11,55	49,35	3145,33	375,59	3148,38	0,80	0,34	0,50

Źródło. opracowanie własne.

Obliczenia przeprowadzono przy użyciu programu komputerowego napisanego w języku Visual Basic for Application, który jest modulem funkcjonalnym w postaci ma-

kra programu Microsoft Excel. Po wprowadzeniu danych w postaci szeregu i wartości wejściowych parametrów wygładzania α, β, γ oraz przyjęciu wartości parametrów $\varepsilon_\alpha, \varepsilon_\beta, \varepsilon_\gamma$, progu λ i wagi ν należy nacisnąć przycisk „Dostrajanie parametrów”. Program dokonuje obliczeń zgodnie z opisaną procedurą i przedstawia obliczone wartości parametrów i kolejne wartości szeregu wygładzonego S_t , oceny przyrostu trendu liniowego T_t , oceny przyrostu trendu kwadratowego R_t , jak również kolejne wartości parametrów wygładzania α, β, γ .

4. Algorytm genetyczny

Algorytm zaproponowany tutaj do zastosowania jest klasycznym algorytmem genetycznym, którego osobnikami są łańcuchy binarne. Działanie algorytmu opiera się na darwinowskiej teorii przetrwania osobników najlepiej przystosowanych. Schemat działania takiego algorytmu przedstawiony jest na rys. 1.

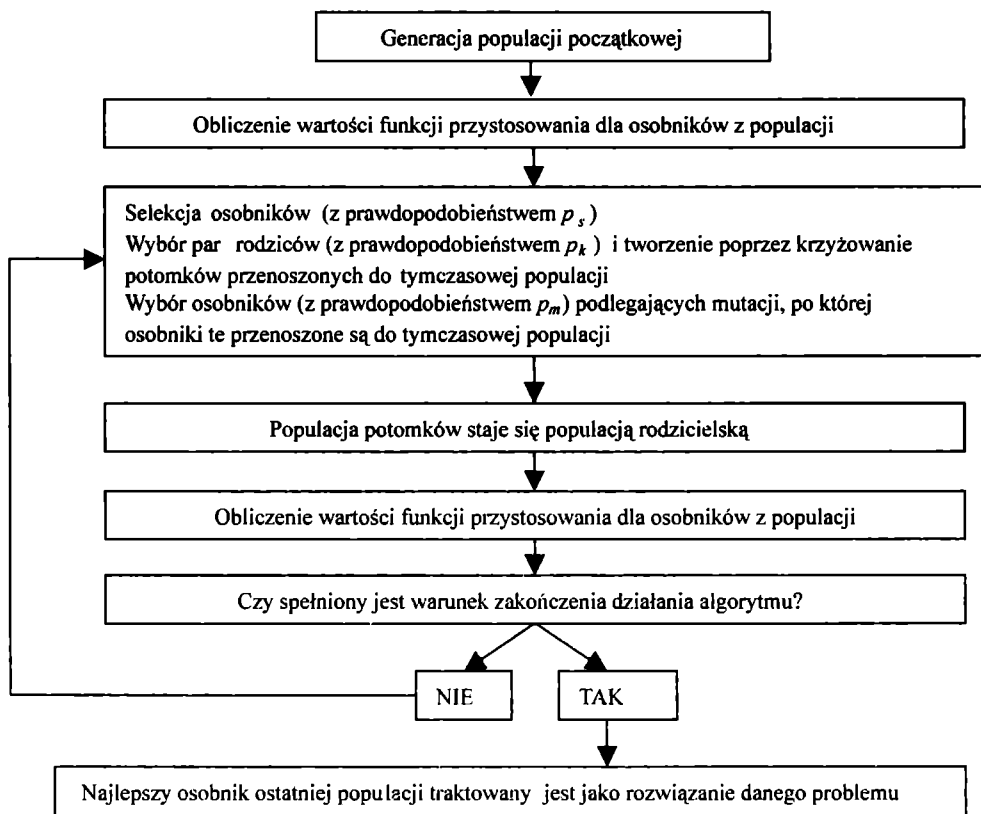
W rozważanym przypadku każdy osobnik, czyli łańcuch składający się z zer i jedynek, reprezentuje zakodowaną wartość trzech parametrów wygładzania α, β, γ . Krzyżowanie jest jednopunktowe, a mutacja polega na zamianie losowo wybranej wartości chromosomu na wartość przeciwną (czyli 0 na 1 lub 1 na 0). Funkcją przystosowania jest błąd średni prognoz wygasłych, przy czym za osobnika lepiej przystosowanego uważa się takiego, dla którego funkcja przystosowania przyjmuje wartość najmniejszą.

Do badań wykorzystano program napisany w języku Visual Basic for Application, który jest modulem funkcjonalnym w postaci makra programu Microsoft Excel. Umieszczenie wyników obliczeń i działania algorytmu genetycznego w arkuszu programu Excel umożliwia przejrzyste przedstawienie otrzymanych wyników i ułatwia ich interpretację.

Parametrami wejściowymi definiowanymi przez użytkownika programu są:

- rozmiar populacji, czyli liczba osobników składających się na każde pokolenie,
- długość chromosomu, która wpływa na dokładność, z jaką wyznaczane są wartości parametrów wygładzania,
- liczba generacji,
- prawdopodobieństwo krzyżowania,
- prawdopodobieństwo mutacji.

Aby rozpocząć działanie procedury, należy wprowadzić szereg czasowy jako kolumnę w pierwszy arkusz, następnie należy wpisać parametry wejściowe algorytmu genetycznego i nacisnąć przycisk „Uruchom program SGA”. Po uruchomieniu algorytmu genetycznego automatycznie przeliczane będą prognozy dla takiego zestawu parametrów wygładzania α, β, γ , dla którego średni błąd prognoz wygasłych jest najmniejszy. W arkuszu uwidocznione są kolejne wartości zarówno szeregu wygładzonego S_t , oceny przyrostu trendu liniowego T_t , oceny przyrostu trendu kwadratowego R_t oraz kolejne wartości parametrów wygładzania α, β, γ .



Rys. 1. Schemat działania algorytmu genetycznego

Źródło: opracowanie własne.

Parametry programu SGA

Rozmiar populacji:	20
Długość chromosomu:	30
Ilość generacji:	30
Prawdopodobieństwo krzyżowania:	0,90
Prawdopodobieństwo mutacji:	0.05
Ograniczenie dolne 1 zmiennej	0,0000
Ograniczenie górne 1 zmiennej	1,0000
Ograniczenie dolne 2 zmiennej	0,0000
Ograniczenie górne 2 zmiennej	1,0000
Ograniczenie dolne 3 zmiennej	0,0000
Ograniczenie górne 3 zmiennej	1,0000

Tabela 2. Przykładowe wyniki obliczeń programu realizującego proponowaną metodę zastosowania AG

t	x_t	s_t	l_t	r_t	$\hat{x}_t$	e_t	α	β	γ
1	3133,5000	3133,5000	-61,5000						
2	3072,0000	3072,0000	-61,5000	84,5000	3052,7500				
3	3095,0000	3095,0000	-8,7185	84,5000	3128,5315	-0,0010	0,6334	0,6246	0,3842
4	3128,5000	3130,2772	-7,6433	12,7714	3129,0196	-0,0004	0,9580	0,0244	0,9932
5	3129,0000	3135,3928	4,9285	12,2258	3146,4342	-111,6352	0,0020	0,9853	0,0127

Źródło: opracowanie własne.

Tabela 3. Przykładowe wartości parametrów uzyskane proponowaną metodą zastosowania AG

Przedział do:	Alfa	Beta	Gamma
0,1	34	75	27
0,2	5	7	5
0,3	7	5	10
0,4	6	6	9
0,5	5	2	9
0,6	6	8	10
0,7	6	10	6
0,8	7	7	8
0,9	12	7	7
1,0	76	37	73
Razem	164	164	164

0,1	20,73%	45,73%	16,46%
0,2	3,05%	4,27%	3,05%
0,3	4,27%	3,05%	6,10%
0,4	3,66%	3,66%	5,49%
0,5	3,05%	1,22%	5,49%
0,6	3,66%	4,88%	6,10%
0,7	3,66%	6,10%	3,66%
0,8	4,27%	4,27%	4,88%
0,9	7,32%	4,27%	4,27%
1,0	46,34%	22,56%	44,51%
Razem	100,00%	100,00%	100,00%

Źródło: opracowanie własne.

Po zakończeniu działania algorytmu genetycznego zostaje wyznaczony rozkład wartości parametrów α , β , γ w przedziałach o rozpiętości 0,1 (tak jak na rys. 3).

Propozycją autorki jest, aby do procedury dostrajania parametrów wybrać te wartości początkowe parametrów wygładzania, które były najczęściej wyznaczane przez algorytm genetyczny.

Działanie takie pozwala na zastosowanie adaptacyjnej procedury dostrajania w sytuacji niedostatecznie długiego szeregu czasowego.

5. Przykład zastosowania

Do badania wzięto szereg czasowy cen miedzi na Londyńskiej Giełdzie Metali z okresu od początku stycznia do końca sierpnia 2005 r. Szereg ten składał się ze 164 obserwacji dziennych.

Tabela 4. Przykładowe wartości parametrów uzyskane proponowaną metodą zastosowania AG

Przedział do:	Alfa	Beta	Gamma
0,1	36	76	37
0,2	2	8	5
0,3	2	6	2
0,4	3	6	7
0,5	8	2	11
0,6	7	7	8
0,7	2	4	6
0,8	10	9	2
0,9	9	4	9
1,0	85	42	77
Razem	164	164	164

0,1	21,95%	46,34%	22,56%
0,2	1,22%	4,88%	3,05%
0,3	1,22%	3,66%	1,22%
0,4	1,83%	3,66%	4,27%
0,5	4,88%	1,22%	6,71%
0,6	4,27%	4,27%	4,88%
0,7	1,22%	2,44%	3,66%
0,8	6,10%	5,49%	1,22%
0,9	5,49%	2,44%	5,49%
1,0	51,83%	25,61%	46,95%
Razem	100,00%	100,00%	100,00%

Źródło: opracowanie własne.

Dla wielokrotnych uruchomień programu realizującego algorytm genetyczny efekty były podobne do tych przedstawionych na rys. 5, tzn. wartości parametrów wygładzania α , β , γ były najczęściej w przedziałach od 0 do 0,1 oraz od 0,9 do 1. Również procedura dostrajania parametrów dawała najmniejszy błąd prognozy przy wartościach początkowych parametrów α , β , γ z tych zakresów. Potwierdza to możliwość wykorzystania algorytmu genetycznego do szacowania początkowych wartości parametrów wygładzania w modelu Holta, które to wartości mogą być później zoptymalizowane przy wykorzystaniu adaptacyjnej procedury dostrajania parametrów.

6. Wnioski

Przeprowadzone badania wskazują na możliwość zastosowania proponowanej metody, konieczne jest jednak przebadanie jej dla innych szeregów czasowych, szczególnie tych wykazujących się zdecydowanie wyraźniejszą sezonowością lub występowaniem trendu nieliniowego.

Literatura

- Dittmann P., *Prognosowanie w przedsiębiorstwie*, Oficyna Ekonomiczna, Kraków 2004.
- Holt C.C., *Forecasting Trends and Seasonal by Exponentially Weighted Moving Averages*, Carnegie Institute of Technology, Pittsburgh, Pennsylvania 1957.
- Klemiato M., Duda J.T., *Mathematical Prediction of Time-Series in Management and Control*, Proceedings of the Summer School on Computer Science in Management, Kraków, May 7-12, 1998.
- Makridakis S., Wheelwright S.C., *Forecasting Methods for Management*, J. Wiley and Sons, New York 1989.
- Prognosowanie gospodarcze. Metody i zastosowanie*, red. M. Cieślak, Wydawnictwo Naukowe PWN, Warszawa 2001.
- Taylor S., *Modelling Financial Time Series*, 1986.
- Zawadzki J., *Ekonometryczne metody prognozowania w przedsiębiorstwie*, Szczecin, Uniwersytet Szczeciński: Rozprawy i Studia, t. 58, 1989.

ABOUT POSSIBLE APPLICATION OF GENETIC ALGORITHMS FOR PREDICTION

Summary

In this article there is shown example of using Holt model prediction with three parameters (α , β , γ) where parameters are estimated by procedure used genetic algorithm. Using this method emerged from the fact that parameters (α , β , γ) should be chosen to minimize mean error of prediction. First of all, the model should follow values changes in time, secondly should distinguish between incidental changes of incoming series and permanent changes of prediction values. Proposed method is based on idea of estimation primary values of pa-

rameters by using genetic algorithm and then optimization with exponential smoothing Holt's formula completed with new adaptive mechanisms.