

## WYKŁADNIKI LAPUNOWA A BŁĄD PROGNOZY

---

### Wstęp

Dokładność wyznaczonej prognozy zależy od własności zbioru danych, takich jak liczba obserwacji  $N$ , poziom szumu  $S$  czy czas ekstrapolacji  $T$ . Błędy prognoz zależą również od charakterystyk systemów dynamicznych: wymiaru atraktora  $D$  oraz spektrum Lapunowa  $\{\lambda_i\}$ . Wykładniki Lapunowa są jednym z narzędzi teorii nieliniowych systemów dynamicznych. Pozwalają na określenie tempa rozchodzenia się sąsiednich trajektorii w przestrzeni stanów. Dodatni wykładnik mierzy wykładniczą rozbieżność, natomiast ujemny – wykładniczą zbieżność dwóch najbliższych trajektorii. Wielkość błędów prognozy zależy również od sposobu prognozowania. Dla prognozowania iteracyjnego błąd rośnie zgodnie z wartością największego wykładnika Lapunowa  $\lambda_{\max}$ , podczas gdy dla prognozowania bezpośredniego –  $q\lambda_{\max}$ .

Celem opracowania jest oszacowanie przyszłych wartości indeksu giełdowego WIG oraz zbadanie zależności pomiędzy błędami prognozy a wykładnikiem Lapunowa. Do wyznaczenia prognozy wykorzystano model autoregresyjny ARCH(m). W opracowaniu wzięto pod uwagę notowania WIG pochodzące z okresu od pierwszego notowania na giełdzie do 21.11.2003 r.

### 1. Wykładniki Lapunowa

Jednym z narzędzi teorii nieliniowych systemów dynamicznych są wykładniki Lapunowa, które mierzą wrażliwość systemów na zmianę warunków początkowych. Określają, jak szybko sąsiednie trajektorie rozchodzą się w przestrzeni stanów. Dodatni wykładnik mierzy rozciąganie się przestrzeni stanów, czyli określa, w jaki sposób prognoza oparta na niedokładnym szacunku danych wyjściowych będzie odbiegać od rzeczywistego rozwoju systemu. Ujemny wykładnik Lapunowa jest miarą zbieżności punktów przestrzeni stanów, czyli określa, ile czasu potrzebuje system na powrót do stanu pierwotnego po wcześniejszym jego zaburzeniu. Dla  $m$ -wymiarowego systemu dynamicznego istnieje  $m$  wykładników Lapunowa.

Dla systemu dynamicznego opisanego za pomocą równania rekurencyjnego:

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

gdzie:

$$f: X \rightarrow X; X \subset R^m,$$

$X$  – przestrzeń stanów,

$$x_t, x_{t+1} \in X,$$

z warunkiem początkowym  $x_0$ , wykładniki Lapunowa są zdefiniowane wzorem:

$$\lambda(x_0) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n \ln |f'(x_t)|, \quad \text{dla } m = 1 \quad (2)$$

oraz uogólniając na przypadek wielowymiarowy:

$$\lambda^i(x_0) = \lim_{n \rightarrow \infty} \frac{1}{n} \ln |\mu_i(n, x_0)|, \quad i = 1, \dots, m, \quad \text{dla } m \geq 1 \quad (3)$$

gdzie  $\mu_i(n, x_0)$  są wartościami własnymi macierzy  $Df^n(x_0)$ ,  $Df^n(x_0)$  jest macierzą Jacobiego odwzorowania  $f^n$  równą:

$$Df^n(x_0) = Df(x_{n-1}) \cdot \dots \cdot Df(x_1) Df(x_0)$$

gdzie  $Df(x) = \left[ \frac{\partial f_i}{\partial x_j}(x) \right]$ ,  $f_j$  są składowymi odwzorowania  $f$ ,  $i, j = 1, 2, \dots, m$ .

## 2. Wykładniki Lapunowa wyższych rzędów

Rozważmy jednowymiarowy system dynamiczny opisany równaniem (1). Wykładnik Lapunowa rzędu  $q$  definiujemy następująco:

$$\lambda^{(q)} = \lim_{t \rightarrow \infty} \frac{1}{t} \log |x_t^{(q)}| \quad (4)$$

gdzie:

$$x_t^{(q)} = \frac{d^q f^t}{dx^q}(x_0)$$

W szczególności dla  $q = 1$  otrzymujemy wykładnik Lapunowa rzędu 1.

Zakładając, że  $f$  jest analityczna oraz że  $f$  jest ergodyczna z naturalną miarą i odpowiednią funkcją gęstości prawdopodobieństwa  $P(x)$ , możemy zapisać wzór (4) w postaci:

$$\lambda^{(q)} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \log |x_i^{(q)}| = \lim_{t \rightarrow \infty} \frac{1}{t} \int \log \left| \frac{d^q f^t}{dx^q} (x) \right| P(x) dx \quad (5)$$

Definiujemy nowy zbiór wykładników:

$$l^{(q)} = \lim_{n \rightarrow \infty} \frac{1}{n} \log \sum_{i=1}^n |x_i^{(q)}| \quad (6)$$

Łatwo można zauważyć, że  $l^{(q)} \geq \lambda^{(q)}$ . W wielu przypadkach prawdziwe jest przybliżenie:

$$l^{(q)} \approx \lambda^{(q)} \quad (7)$$

Przybliżenie (7) można wykorzystać do pokazania związku pomiędzy wykładnikami Lapunowa pierwszego rzędu a wykładnikami Lapunowa wyższych rzędów [3]:

$$\lambda^{(q)} \approx q \lambda^{(1)} \quad (8)$$

### 3. Prognozowanie bezpośrednie

Prognozowanie bezpośrednie polega na wyznaczeniu dla każdej prognozowanej wartości nowego przybliżenia  $\hat{f}_T$ , tj. zbudowania nowego modelu dla każdego  $T$ . Załóżmy, że wykorzystujemy liniową interpolację i przybliżamy funkcję  $F$  na przedziale  $[x_1, x_2]$ :

$$F(x) \approx \hat{F}(x) = F(x_1) + \frac{F(x_2) - F(x_1)}{x_2 - x_1} (x - x_1) \quad (9)$$

Błąd powyższej aproksymacji wynosi  $E(x) = F(x) - \hat{F}(x)$ . Rozwijając  $F(x)$  w szereg Taylora względem  $x_1$ , otrzymujemy:

$$E(x) = \frac{F''(x_1)}{2} (x - x_1)(x - x_2) + O(\varepsilon^3) \quad (10)$$

gdzie  $\varepsilon = x_2 - x_1$ .

$O(\varepsilon^n)$  wskazuje, że pozostałe składniki są rzędu co najwyżej  $\varepsilon^n$ .  
Średni absolutny błąd jest równy:

$$\langle |E(x)| \rangle = \frac{1}{\varepsilon} \int_{x_1}^{x_2} |E(x)| dx = \frac{1}{12} |F''(x_1)| \varepsilon^2 + O(\varepsilon^3) \quad (11)$$

Rozważmy jednowymiarowe odwzorowanie:

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

Założmy, że odwzorowanie  $f$  jest ergodyczne. Przybliżając iterację  $T$  oraz wykorzystując liniową interpolację, otrzymamy:

$$\langle |E(x_{t+T})| \rangle = \frac{\varepsilon^2}{12} \int \left| \frac{d^2 f^T}{dx^2}(x) \right| P(x) dx + O(\varepsilon^3) = \frac{\varepsilon^2}{12} |x_T''| + O(\varepsilon^3) \quad (12)$$

Uwzględniając równanie (6) oraz (7), otrzymujemy:

$$\log \langle |E(x_{t+T})| \rangle \sim q t^{(1)} T \approx q \lambda^{(1)} T \quad (13)$$

Zatem średni błąd prognozy:

$$\log \langle |E(x_{t+T})| \rangle \sim q \lambda^{(1)} T \quad (14)$$

## 4. Prognozowanie iteracyjne

Prognozowanie iteracyjne polega na zbudowaniu nowego modelu dla każdego  $T$ . Założmy, że  $T=1$  oraz że przybliżamy funkcję  $f$  za pomocą funkcji  $g$ . Wówczas błąd przybliżenia wynosi:

$$\delta(x) = g(x) - f(x) \quad (15)$$

Podobnie dla dowolnego  $T$  błąd przybliżenia jest równy:

$$\Delta_T(x) = g^T(x) - f^T(x) \quad (16)$$

gdzie  $\Delta_1 = \delta$ .

Zakładamy również, że dla dowolnego  $T$  nasza aproksymacja jest wystarczająco dobra, taka że błąd  $\Delta_T(x)$  jest mały i ograniczony nierównościami:

$$\Delta_{\max} > |\Delta_T| \quad (17)$$

dla każdego  $x$ .

Z równań (15) i (16) otrzymujemy:

$$\Delta_T(x) = f(g^{T-1}(x)) + \delta(g^{T-1}(x)) - f^T(x) \quad (18)$$

Rozwijamy  $f(g^{T-1}(x))$  oraz  $\delta(g^{T-1}(x))$  w szereg Taylora pierwszego rzędu. Zakładając, że  $\delta$  oraz  $f$  są gładkie, takie że  $f'' = O(1)$  oraz  $\delta'' = O(1)$ , otrzymujemy:

$$\Delta_T(x) = g'(f^{T-1}(x))\Delta_{T-1}(x) + \delta(f^{T-1}(x)) + O(\Delta_{\max}^2) \quad (19)$$

Rozszerzając wyrażenie (19) dla  $T-2, T-3, \dots$ , otrzymamy:

$$\Delta_T(x) = \sum_{j=0}^{T-1} \prod_{i=j+1}^{T-1} g'(f^i(x))\delta(f^i(x)) + O(\Delta_{\max}^2) \quad (20)$$

Przyjmując założenie, że  $\delta' = O(\Delta_{\max})$ , możemy przybliżyć  $g'$  za pomocą  $f'$ . Korzystając ze wzoru na pochodną funkcji złożonej, możemy zapisać:

$$\Delta_T(x) = \sum_{j=1}^T \frac{df^{T-j}}{dx}(x_j)\delta(x_{j-1}) + O(\Delta_{\max}^2) \quad (21)$$

Biorąc wartość bezwzględną i uśredniając po  $x$ , mamy:

$$\langle |\Delta_T(x)| \rangle = \sum_{j=1}^T \left\langle \left| \frac{df^{T-j}}{dx}(x_j)\delta(x_{j-1}) \right| \right\rangle + O(\Delta_{\max}^2) \quad (22)$$

Dla uproszczenia założmy, że  $f(x)$  jest jednowymiarowym odwzorowaniem oraz że  $\frac{df^{T-j}}{dx}(x_j)$  i  $\delta(x_{j-1})$  są nieskorelowane.  $\langle |\delta(x_j)| \rangle$  nie zależy od  $j$ . Uwzględniając definicję (6), średnia pochodnych  $\left\langle \left| \frac{df^j}{dx}(x_0) \right| \right\rangle \approx L^j$ , gdzie  $L = e^{l^{(1)}}$ . Wówczas możemy zapisać:

$$\langle |\Delta_T(x)| \rangle \approx \langle |\delta(x)| \rangle \sum_{i=0}^{T-1} L^i \quad (23)$$

Sumując ten szereg, otrzymujemy:

$$\langle |\Delta_T(x)| \rangle \approx \frac{L^T - 1}{L - 1} \langle |\delta(x)| \rangle \quad (24)$$

Dla dużych  $T$  mamy:

$$\lim_{T \rightarrow \infty} \left( \frac{L^T - 1}{L - 1} \right)^{\frac{1}{T}} = L \quad (25)$$

Ponownie przybliżając  $g$  za pomocą  $f$  oraz powracając z podstawieniem dla  $\Delta_{T-i}$ , otrzymujemy  $\Delta_T(x) = \frac{df^T}{dx}(x)\delta(x)$ . Logarytmując i biorąc wartość bezwzględną oraz zakładając, że  $\delta(x)$  jest nieskorelowane z  $\frac{df^T}{dx}(x)$ , otrzymujemy:

$$\langle \log|\Delta_T(x)| \rangle = \left\langle \log \left| \frac{df^T}{dx}(x) \right| \right\rangle + \langle \log|\delta(x)| \rangle = T\lambda^{(1)} + \langle \log|\delta(x)| \rangle \quad (26)$$

## 5. Algorytm prognozowania

Do oszacowania przyszłych wartości szeregu czasowego wykorzystano model o charakterze autoregresyjnym. W celu wyznaczenia prognozy konieczna jest jedynie znajomość wartości zmiennej endogenicznej w poprzednich okresach.

Model ARCH( $m$ ) można zapisać w następującej postaci:

$$Z_t = \text{const} + \sqrt{h_t} \varepsilon_t \quad (27)$$

$$h_t = \alpha_0 + \sum_{j=1}^m \alpha_j Z_{t-j}^2, \quad t = 1, \dots, n \quad (28)$$

gdzie:

$Z_t$  – logarytmiczna stopa zwrotu,

$h_t$  – funkcja wariancji warunkowej,

$\alpha_0 > 0, \alpha_j \geq 0$  – parametry estymowane metodą największej wiarygodności,  
 $j = 1, \dots, m,$

$\varepsilon_t$  – składnik losowy.

Parametry funkcji warunkowej wariancji oszacowano za pomocą programu GRETL [5].

Do oceny dokładności prognozy wykorzystano względny błąd prognozy w momencie  $t$ :

$$D_t = \left| \frac{x_t - \hat{x}_t}{x_t} \right| \cdot 100\% \quad (29)$$

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 absolutny błąd prognozy dany wzorem:

$$E = \frac{1}{T} \sum_{t=1}^T \left| \frac{\hat{x}_t - x_t}{x_t} \right|, \quad T = 1, 2, \dots, h \quad (30)$$

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.

Tabele 1 i 2 zawierają wyniki prognozowania iteracyjnego oraz bezpośredniego indeksu giełdowego WIG.

Tabela 1

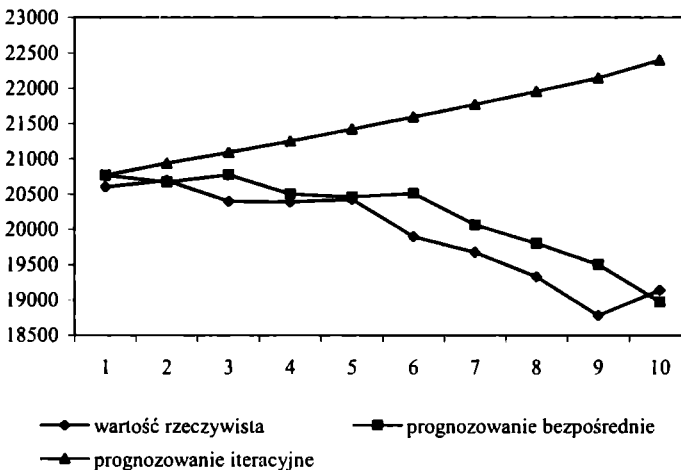
Wyniki prognozowania iteracyjnego indeksu giełdowego WIG

$T$	$x_t$	$\hat{x}_t$	$D_t$	$E$
1	20602,63	20770,33	0,814%	0,814%
2	20691,85	20671,97	0,0961%	0,455%
3	20401,72	20774,58	1,8276%	0,9125%
4	20388,22	20504,61	0,5709%	0,8271%
5	20425,08	20464,02	0,1906%	0,6998%
6	19899,74	20510,54	3,0694%	1,0948%
7	19674,12	20065,92	1,9915%	1,2229%
8	19333,5	19802,87	2,4278%	1,3735%
9	18782,83	19504,97	3,8447%	1,6481%
10	19144,28	18973,44	0,8924%	1,5725%

Tabela 2

## Wyniki prognozowania bezpośredniego indeksu giełdowego WIG

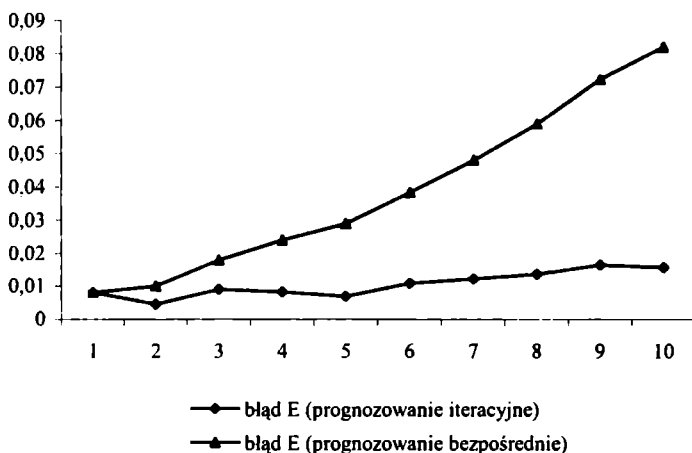
T	$x_t$	$\hat{x}_t$	$D_t$	E
1	20602,63	20770,33	0,814%	0,814%
2	20691,85	20938,91	1,194%	1,004%
3	20401,72	21092,38	3,3853%	1,7977%
4	20388,22	21250,21	4,2279%	2,4053%
5	20425,08	21420,51	4,8735%	2,8989%
6	19899,74	21593,31	8,5105%	3,83342%
7	19674,12	21771,08	10,6585%	4,8091%
8	19333,5	21955,6	13,5625%	5,9033%
9	18782,83	22145,3	17,9018%	7,2364%
10	19144,28	22397,63	16,9939%	8,2122%



Rys. 1. Wyniki prognozowania przyszłych wartości indeksu giełdowego WIG

Na rys. 2 przedstawiono średnie absolutne błędy otrzymanych prognoz metodą iteracyjną oraz bezpośrednią.





Rys. 2. Średni absolutny błąd prognozy

Do wyznaczenia największego wykładnika Lapunowa posłużono się programem opartym na algorytmie Wolfa, zamieszczonym jako dodatek w pracy [7]. 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 za bardzo się oddalają, znajduje się punkt zastępczy po to, by pominąć etap kurczenia się systemu, ponieważ wykładnik Lapunowa mierzy oddalanie się punktów w przestrzeni stanów, a nie ich zbliżanie.

Wartość największego wykładnika Lapunowa dla indeksu WIG wynosi 0,00001327.

## Podsumowanie

W opracowaniu wyznaczono przyszłe wartości szeregu czasowego utworzonego z indeksu giełdowego WIG. Tabele 1 i 2 zawierają prognozy uzyskane za pomocą prognozowania odpowiednio metodą iteracyjną i metodą bezpośrednią. Na rys. 2 zaznaczono średnie absolutne błędy prognozy dla poszczególnych sposobów prognozowania. Analiza otrzymanych wyników pokazuje, że prognozy otrzymane metodą iteracyjną są dokładniejsze. Świadczą

o tym niższe wartości otrzymanych prognoz. W przypadku prognozowania bezpośredniego, średni błąd absolutny prognozy jest proporcjonalny do wartości największego wykładnika Lapunowa.

## 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. Kufel T.: *Ekonometria. Rozwiązywanie problemów z wykorzystaniem programu GRETL*. Wydawnictwo Naukowe PWN, Warszawa 2004.
6. Miśkiewicz M.: *Wykładniki Lapunowa na Gieldzie Papierów Wartościowych w Warszawie*. Zeszyty Naukowe AE w Katowicach (w druku).
7. Peters E.E.: *Teoria chaosu a rynki kapitałowe. Nowe spojrzenie na cykle, ceny i ryzyko*. WIG-Press, Warszawa 1997.
8. Wolf A., Swift J.B., Vastano J.A.: *Determining Lyapunov Exponents from a Time Series*. „Physica 16D” 1985, s. 285-317.

## LYAPUNOV EXPONENTS AND THE FORECASTING ERRORS

### Summary

In this paper the forecasting errors have been estimated. The errors depend on the properties of the dynamics, such as the attractor dimension and the Lyapunov spectrum. The Lyapunov exponents of a dynamic system are measures of the degree of the divergence of the nearby orbits in a state space. The rate at which errors grow depends on the way we make predictions: iterative forecasts by fitting a model for  $T = 1$  and iterating to make predictions for  $T = 2, 3, \dots$ , or direct forecasts by fitting a new model for each individual  $T$ . The errors for iterative forecasts grow at an exponential rate given by  $\lambda_{\max}$ , and for direct forecasts  $q\lambda_{\max}$ .