PRACE NAUKOWE AKADEMII EKONOMICZNEJ WE WROCŁAWIU
Nr 1088 — 2005
Inwestycje finansowe i ubezpieczenia – tendencje światowe a polski rynek

Ewa Broszkiewicz-Suwaj, Agnieszka Wyłomańska
Politechnika Wrocławska

OKRESOWA KORELACJA A INTEGRACJA I KOINTEGRACJA

1. Wstęp

W artykule pokazano relację pomiędzy okresową korelacją a integracją szeregu czasowego. Jest to temat mający ważne znaczenie w analizie szeregów czasowych, a więc znajdujący zastosowanie w analizie danych dotyczących rynków kapitałowych.

Autorki podjęły temat badania szczególnego rodzaju szeregu czasowego – szeregu, który ma okresową funkcję korelacji (szeregu okresowo skorelowanego). Tradycyjne podejście do badania takich szeregów bazuje w znacznej mierze na analizie spektralnej i najczęściej modeluje się je szeregami PARMA (*periodic* ARMA). Przy estymacji współczynników takich modeli występują jednak duże błędy ze względu na znaczną liczbę estymowanych parametrów. Dlatego podjęto próbę innego podejścia do szeregów okresowo skorelowanych. Autorki dowodzą, że szeregi te można podzielić na zintegrowane podszeregi. Istnieją zatem również podstawy do badania ich kointegracji. Podejście takie może mieć kluczowe znaczenie w analizie danych wykazujących różnego rodzaju okresowości. Jest to nowe i oryginalne ujęcie, wiążące się z ciekawymi badaniami, które mogą dać interesujące wyniki.

W pierwszej części pracy omówiono własności procesów okresowo skorelowanych oraz przedstawiono jedną z metod opartą na analizie spektralnej, służącą do wykrywania okresu funkcji korelacji. Metodę tę wykorzystano do zbadania okresowej korelacji w danych rzeczywistych opisujących średnie dzienne spotowe ceny energii elektrycznej na Giełdzie Nord Pool z okresu od 30 grudnia 1996 r. do 1 grudnia 1998 r. W drugim punkcie zajęto się zagadnieniem integracji oraz zbadano ją dla podszeregów wyznaczonych z szeregu wyjściowego. Podszeregi te opisują spotowe ceny energii w poszczególnych dniach tygodnia. Na zakończenie omówiono zjawisko kointegracji i przetestowano hipotezę o jej istnieniu.

2. Procesy okresowo skorelowane

Analiza szeregów czasowych w znacznej mierze bazuje na stacjonarności danych. Jednak w większości przypadków założenie, że badany szereg jest stacjonarny (to znaczy jego średnia i kowariancja są niezmiennicze na przesunięcia [Hamilton 1994]) jest uproszczeniem i może prowadzić do błędnych analiz. Procesy okresowo skorelowane są klasą procesów, które generalnie są niestacjonarne, ale wykazują wiele własności procesów stacjonarnych. Spotykane są m.in. w klimatologii, hydrologii i ekonomii.

Definicja 1. Proces stochastyczny $\{X(t),\ t \in I\}$ o skończonym drugim momencie nazywany jest okresowo skorelowanym (ang. *periodically correlated*) z okresem T, jeśli T jest najmniejszą liczbą taką, że dla każdych t oraz u ze zbioru I spełnione są następujące warunki [Gladyshev 1961]:

(*i*) $m_X(t) = E(X(t)) = m_X(t+T)$,

(*ii*) $R_X(t,u) = E((X(t) - m_X(t))(X(u) - m_X(u))) = R_X(t+T, u+T)$.

Szereg czasowy okresowo skorelowany jest stacjonarny dla $T = 1$.

Poniżej przedstawiono algorytm wykrywania okresowej korelacji dla zadanego szeregu czasowego $\{X(1), X(2), ..., X(N)\}$ o średniej $m_X(n) \equiv 0$ (n = 1, 2, ...). Jest to metoda oparta na analizie spektralnej [Broszkiewicz-Suwaj i in. 2004; Hurd, Gerr 1991].

Schemat algorytmu

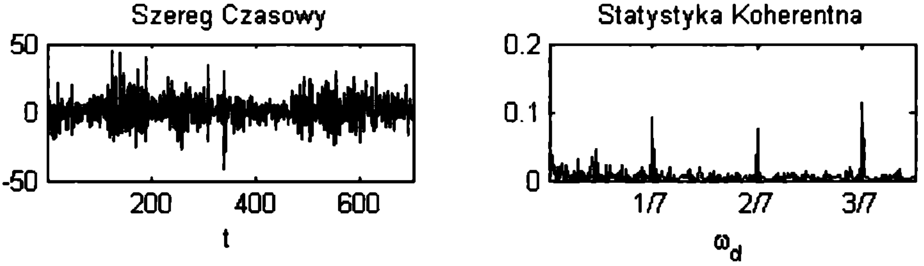
1. Wyznaczamy dyskretną transformatę Fouriera $I_N(\omega) = \sum_{n=1}^{N} X(n)\exp(-i\omega(n-1))$ w punktach $\omega_k = \dfrac{2\pi(k-1)}{N}$, $k = 1, 2, \ldots, N$.

2. Obliczamy dwuwymiarową statystykę zwaną koherencją próbkową (ang. *sample coherence)*

$$|\gamma(\omega_p, \omega_q, M)|^2 = \frac{\left|\sum_{m=0}^{M-1} I_N(\omega_{p+m})\overline{I_N(\omega_{q+m})}\right|^2}{\sum_{m=0}^{M-1}|I_N(\omega_{p+m})|^2 \sum_{m=0}^{M-1}|I_N(\omega_{q+m})|^2}$$ w każdym punkcie (p,q).

3. Wyznaczamy wartości jednowymiarowej statystyki koherentnej (ang. *coherent statistic*) $|\gamma(0, \omega_d, N)|^2$. Statystyka ta przyjmuje wartości z przedziału $[0,1]$. Jeżeli na wykresie pojawią się wyraźne „piki" w punktach $\omega_d, 2\omega_d, 3\omega_d, \ldots$, to znaczy, iż zadany szereg jest okresowo skorelowany z okresem $T = \dfrac{1}{\omega_d}$.

Istnieją również inne metody wykrywania okresowej korelacji oparte na analizie spektralnej [Broszkiewicz-Suwaj 2003: Broszkiewicz-Suwaj i in. 2004; Hurd, Gerr 1991]. Na potrzeby artykułu zbadano okresową korelację danych pochodzących ze skandynawskiej giełdy energii (średnie dzienne spotowe ceny z okresu od 30 grudnia 1996 r. do 1 grudnia 1998 r.). Wybrano tutaj dane z rynku energii elektrycznej ze względu na fakt, że wykazują one okresowość. Poprzez zróżnicowanie danych z siedmiodniowym opóźnieniem przeprowadzono opisany test.



Rys. 1. Szereg czasowy przedstawiający dzienne zróżnicowane tygodniowo ceny z Giełdy Nord Pool oraz wykres statystyki koherentnej dla tego szeregu
Źródło: opracowanie własne.

„Piki" występujące w częstotliwościach 1/7, 2/7,... wskazują na okres funkcji korelacji, który jest równy 7 dni (tygodniowa okresowa korelacja).

3. Integracja i kointegracja

W tej części pracy zostaną omówione zagadnienia integracji i kointegracji oraz metodologia związana z testowaniem tych zjawisk w danym szeregu czasowym. Ponadto zostanie pokazany związek pomiędzy okresową korelacją a integracją.

Definicja 2. Szereg niestacjonarny $\{X(n), n \in Z\}$, który można sprowadzić do szeregu stacjonarnego, obliczając przyrosty $(\Delta X(n) = X(n) - X(n-1))$ d razy, nazywamy szeregiem zintegrowanym stopnia d i oznacza się go $X(n) \sim I(d)$ [Charemza, Deadman 1997; Engle, Granger 1987; Welfe 1998].

Istnieją testy, które pozwalają na zbadanie stopnia integracji, np. test Dickeya-Fullera (DF) oraz rozszerzony test Dickeya-Fullera (ADF) [Charemza, Deadman 1997, Gladyshev 1961]. Do prostego testowania, czy szereg jest zintegrowany, można zastosować test oparty na statystyce Durbina-Watsona:

$$IDW = \frac{\sum (X(n) - X(n-1))^2}{\sum (X(n) - \bar{X}(n))^2}. \quad (1)$$

Jeśli wartość statystyki IDW jest mniejsza niż 0,5, istnieje podejrzenie braku integracji, natomiast gdy jej wartość jest bliska 2, wówczas można przyjąć hipotezę o zintegrowaniu szeregu [Charemza, Deadman 1997].

Uwaga 1. Okresowa korelacja danego szeregu niestacjonarnego $X(n)$, w którym wykryto okresową korelację, jest równoważna integracji $I(0)$ jego podszeregów, powstałych w następujący sposób:

$$Y(v,n) = X(nT+v) \qquad v = 1, 2, ..., T. \tag{2}$$

Dowód

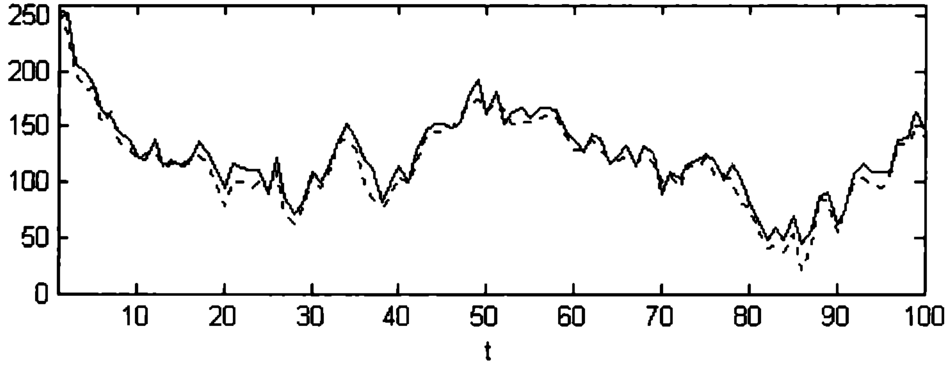
Własność ta wynika z następujących zależności, które spełnione są dla dowolnych n i k ze zbioru liczb naturalnych:

$$m_{Y(v)}(n) = m_X(nT+v) = m_X(kT+v) = m_{Y(v)}(k) = const$$

$$R_{Y(v)}(n,k) = R_X(nT+v, kT+v) = R_X(v, |k-n|T+v) = R_{Y(v)}(0, |k-n|).$$

Zatem dla każdego v = 1, 2, ...T szereg $\{Y(v,n)\}$ jest stacjonarny.

Dla rzeczywistych danych okresową korelację uzyskujemy dopiero po zróżnicowaniu zadanego szeregu, co może spowodować utratę długookresowych zależności. Interesujące może być zatem zachowanie podszeregów powstałych jak we wzorze (2) dla niezróżnicowanego szeregu pierwotnego.



Rys. 2. Dwa przykładowe (piątek, sobota) podszeregi powstałe z szeregu opisującego ceny energii na Giełdzie Nord Pool
Źródło: opracowanie własne.

Poniżej przedstawiono wyniki otrzymane przy badaniu integracji dla podszeregów opisujących ceny energii z Giełdy Nord Pool w poszczególnych dniach tygodnia (7 podszeregów). Do badania stopnia integracji wykorzystano szybką metodę badania hipotezy o kointegracji opartą na statystyce Durbina-Watsona oraz wykresy funkcji autokorelacji (ACF).

Tabela 1. Wartość statystyki *IDW* dla szeregów obejmujących ceny w poszczególnych dniach tygodnia

Cena	Poniedziałek	Wtorek	Środa	Czwartek	Piątek	Sobota	Niedziela
IDW	0,2187	0,1819	0,2470	0,2606	0,2034	0,1507	0,1955

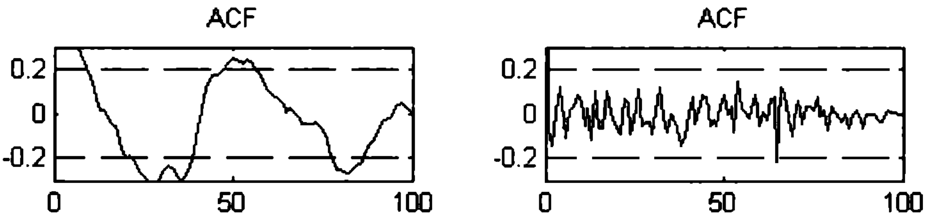
Źródło: opracowanie własne.

Wartość statystyki *IDW* dla danych pierwotnych (tab. 1) jest istotnie mniejsza niż 0,5 dla każdego z szeregów, zatem możemy odrzucić hipotezę o integracji stopnia 0 dla każdego z nich. Wartości statystyki *IDW* dla przyrostów pierwszego rzędu poszczególnych szeregów (tab. 2) wskazują na fakt, że szeregi te są stacjonarne, zatem można wnioskować o integracji $I(1)$ poszczególnych podszeregów.

Tabela 2. Wartość statystyki *IDW* dla przyrostów szeregów obejmujących ceny w poszczególnych dniach tygodnia

Cena	Poniedziałek	Wtorek	Środa	Czwartek	Piątek	Sobota	Niedziela
IDW	2,1190	1,9728	2,1501	2,3200	2,2453	1,8699	2,1508

Źródło: opracowanie własne.



Rys. 3. Przykładowy wykres funkcji autokorelacji dla podszeregu w wersji niezróżnicowanej (panel lewy) i zróżnicowanej (panel prawy). Przerywana linia oznacza 95% przedziały ufności dla ruchu Browna
Źródło: opracowanie własne.

Dla wszystkich badanych podszeregów wykresy funkcji autokorelacji mają zbliżoną postać. Podobnie jak poprzedni test, wskazują one na brak stacjonarności przed różnicowaniem i jej obecność po zróżnicowaniu.

Z pojęciem integracji jest ściśle związane zagadnienie kointegracji.

Definicja 3 [Hamilton 1994]. Mówimy, że szeregi czasowe $Y(1,n)$, $Y(2,n)$, ...,$Y(m,n)$ są skointegrowane stopnia d, b, gdzie $d \geq b \geq 0$ i piszemy $Y(1,n)$, $Y(2,n)$, ...,$Y(m,n) \sim CI\,(d, b)$, jeżeli:

1) są one zintegrowane stopnia d,

2) istnieje kombinacja liniowa tych zmiennych, np. $a_1Y(1,n) + a_2Y(2,n) + ... + a_mY(m,n)$, która jest zintegrowana stopnia $d - b$. Wektor $[a_1, a_2, ..., a_m]$ nazywamy kointegrującym.

Test kointegracji przebiega dwuetapowo. W pierwszym etapie badamy stopień zintegrowania poszczególnych szeregów, co daje wskazówki do dalszego badania kointegracji. W drugim etapie należy wyestymować wektor kointegrujący mający postać $[1, -a_1, -a_2, \ldots, -a_6]$, który pojawia się w następującym modelu regresji:

$$Y(1,n) = a_1Y(2,n) + a_2Y(3,n) + ...a_6Y(7,n) + v(n). \tag{3}$$

Można go estymować, używając metody najmniejszych kwadratów, a następnie testować hipotezę o kointegracji z użyciem testów Dickeya-Fullera. Inną prostą metodą testowania tej hipotezy jest użycie analogonu statystyki Durbina-Watsona. Test ten polega na wyznaczeniu statystyki Durbina-Watsona w celu oszacowania odchyleń szeregu od głównej trajektorii, które przy założeniu prawdziwości hipotezy o kointegracji są stacjonarne [Charemza, Deadman 1997]

$$CIDW = \frac{\sum(\hat{v}(n) - \hat{v}(n-1))^2}{\sum(\hat{v}(n) - \overline{v}(n))^2}, \tag{4}$$

gdzie $\hat{v}(n)$ są resztami wyznaczonymi metodą najmniejszych kwadratów dla równania (3). Jeśli wartość statystyki *CIDW* jest mniejsza od współczynnika determinancji R^2

$$R^2 = 1 - \frac{\sum \hat{v}^2(n)}{\sum (Y(1,n) - \overline{Y}(1,n))^2},$$

postuluje się odrzucenie hipotezy o kointegracji. W przeciwnym wypadku, gdy $CIDW > R^2$, kointegracja może występować [Charemza, Deadman 1997].

Tabela 3. Wartości współczynników kointegracji

*a*1	*a*2	*a*3	*a*4	*a*5	*a*6
1,3473	–0,3781	0,0246	–0,0376	0,2434	–0,2087

Źródło: opracowanie własne.

W analizach badamy kointegrację szeregów, które opisują cenę energii elektrycznej w poszczególnych dniach tygodnia. Za zmienną opisywaną przyjmujemy ceny z poniedziałku; pozostałe szeregi tworzą zmienne opisujące. Dla takiego przypadku współczynnik determinancji wynosi 0,9524, a wartość statystyki *CIDW* jest równa 2,1523, zatem szeregi te są skointegrowane CI(1,1). Współczynniki kointegracji dla tego modelu umieszczone są w tab. 3.

4. Podsumowanie

Przeprowadzona analiza wskazuje, że okresowa korelacja jest ściśle związana z zagadnieniem integracji, a co za tym idzie – z kointegracją. Szereg okresowo skorelowany można przetransformować na zbiór podszeregów, które wykazują własność integracji. Ponadto podszeregi w swej niezróżnicowanej postaci mogą być ze sobą skointegrowane. Należy podkreślić, że sama okresowość szeregu nie wystarcza do tego, by był on złożeniem podszeregów zintegrowanych. W pracy analizowano szereg okresowo skorelowany opisujący spotowe ceny energii elektrycznej. Przedstawiona metodologia pozwala na badanie i opisywanie procesów wykazujących okresową korelację bez utraty zależności długookresowych. Służy ona także jako potwierdzenie słuszności metod wykrywania okresowej korelacji opartych na analizie spektralnej. Ma ona również ważne znaczenie w analizie szeregów czasowych, które wykazują okresowości, a oryginalne podejście do tematu może mieć istotny wpływ na dalszy rozwój badań szeregów okresowo skorelowanych.

Literatura

Broszkiewicz-Suwaj E., *Wykrywanie okresowej korelacji danych z TGE SA w oparciu o analizę spektralną*, „Rynek Terminowy" 2003, nr 20, s. 92-95.

Broszkiewicz-Suwaj E., Makagon A., Weron R., Wyłomańska A., *On Detecting and Modeling Periodic Correlation in Financial Data*, „Physica A" 2004, 336, s. 196-205.

Charemza W.W., Deadman D.F., *Nowa ekonometria*, Polskie Wydawnictwo Ekonomiczne, Warszawa 1997.

Engle R.F., Granger W.J., *Co-Integration and Error Correction: Representation, Estimation, and Testing*, „Econometrica" 1987, 55 (nr 2), s. 251-276.

Gladyshev E.G., *Periodically Correlated Random Sequences*, „Soviet Math." 1961, 2, 385-388.

Hamilton J.D., *Time Series Analysis*, Princeton University Press, Princeton 1994.

Hurd H., Gerr N.L., *Graphical Methods for Detemining the Presence of Periodic Correlation*, „J. Time Ser. Anal." 1991. 12 (4), s. 337-350.

Welfe A., *Ekonometria*, Polskie Wydawnictwo Ekonomiczne, Warszawa 1998.

PERIODIC CORRELATION – INTEGRATION AND COINTEGRATION

Summary

In this paper we present a new approach to integration and cointegration. We show that a periodically correlated time series can be divided in a natural way into subseries that are integrated. Moreover, with high probability they are cointegrated. Therefore it is enough to show periodic correlation of the original series to conclude that the subseries are integrated.

In the first part of the paper we present the main features of periodically correlated processes and a method of detecting periodic correlation. We illustrate it using a data set of spot electricity prices from the Nord Pool Power Exchange. In the next section we show that the subseries (one for each day of the week) exhibit integration as well as cointegration.