

Sylwester Smolik

Katedra Zastosowań Matematyki SGGW

Piotr Łukasiewicz

Katedra Ekonometrii i Informatyki SGGW

Opis zjawisk okresowych z monotonicznie zmieniającą się amplitudą

Wprowadzenie

W badaniach ekonomicznych, technicznych i przyrodniczych bardzo często uzyskujemy wartości badanej cechy czy zjawiska w różnych momentach czasu. Tego rodzaju dane, uporządkowane w ciągu (t, y_t) , dla $t = 1, 2, \dots, n$, nazywamy szeregiem czasowym. Takie zbiory danych prezentujemy za pomocą tabeli lub wykresu, jako punkty płaszczyzny połączone odcinkami linii prostej. Już z wizualnej obserwacji wspomnianego wykresu często wnioskujemy, że badane zjawisko podlega typowym prawidłowościom, których wykrycie i opis ilościowy są celem analizy szeregów czasowych. Patrząc na rysunek 1, obrazujący skup mleka w Polsce w latach 1970–1979, spostrzegamy, że prezentowane zjawisko podlega wyraźnej tendencji wzrostowej (trend) oraz że w kolejnych latach mają miejsce podobne kształtem wahania okresowe, ale o rosnącej z latami amplitudzie. Są one obciążone pewnym błędem pomiaru. W teorii ekonomicznych szeregów czasowych nazywa się je wahaniami przypadkowymi. Powszechnie przyjęta metoda analizy szeregów czasowych polega na rozłożeniu ich wartości na trzy składniki – trend, składnik okresowy (w szczególności sezonowy) i składnik losowy. Postępowanie to tłumaczy się tym, że wymienione wyżej składniki szeregu czasowego są wynikiem działania trzech wyłączających się kompleksów przyczyn kształtujących ich nasilenie. Zgodnie z tym będziemy analizowali takie zjawiska, które można rozłożyć na wymienione wyżej trzy składniki:

$$y_t = f(t) + z(t) + \varepsilon_t \quad (t = 1, 2, \dots, n). \quad (1)$$

Wyodrębnienie w badanym zjawisku tendencji rozwojowej przez eliminację jego wahań okresowych i przypadkowych nazywamy wyrównywaniem lub wygładzaniem szeregu czasowego. Zwykle wyróżnia się dwie metody wygładzania: mechaniczną i analityczną. W mechanicznej metodzie używa się średnich ruchomych. Liczba składników średniej ruchomej powinna być odpowiednio dobrana. W analitycznej metodzie wygładzania tendencję rozwojową badanego zjawiska opisujemy funkcją gładką, a jej parametry estymujemy metodą najmniejszych kwadratów. Naszym zdaniem, parametry funkcji trendu przy obecności wahań okresowych wygodnie jest estymować metodą średnich, analogicznie jak to czyni się w metodzie mechanicznego wygładzania. W ten sposób nie tylko zachowujemy taką samą metodę przy wyznaczaniu tendencji rozwojowej, ale tylko ta metoda zapewnia sprowadzenie oscylacji przekształconego szeregu czasowego wokół zera. Pozwala to na zmniejszenie o jeden liczby parametrów koniecznych do opisu okresowości zjawiska. Dokładniejsze objaśnienia przytoczymy w czwartej części niniejszego artykułu.

Estymacja parametrów modelu wahań okresowych

Zakładamy, że dany jest ciąg punktów empirycznych (t, y_t) dla $t = 1, 2, \dots, n$, obrazujących przebieg badanego zjawiska. Ocenę tendencji rozwojowej $f(t)$ można wyznaczyć różnymi metodami. Wygodnie jest estymować parametry funkcji $f(t)$ wspomnianą wcześniej metodą średnich, wówczas, zgodnie z jej definicją, obliczamy szukane wartości ocen parametrów regresji z warunku zerowania się sumy odchyłeń. Pozostały składnik danego szeregu czasowego $z_t = y_t - \hat{f}(t) - \varepsilon_t$ opiszemy harmoniką postaci:

$$z_t = (A + B \cdot t) \cdot \sin(\omega t + \theta) - \varepsilon_t, \text{ gdzie } \omega = 2\pi/T \quad (t = 1, 2, \dots, n), \quad (2)$$

gdzie: ω – częstość (prędkość kątowna), T – okres badanego zjawiska, θ – jego faza początkowa, A i B to poszukiwane stałe, z których buduje się amplitudę harmoniki. Może być ona stała lub zmieniać się monotonicznie. Parametry modelu (2) szacujemy metodą najmniejszych kwadratów. W tym celu poszukujemy minimum następującej funkcji:

$$F(A, B, \theta, \omega) = \sum_{t=1}^n \varepsilon_t^2 = \sum_{t=1}^n [z_t - (A + B \cdot t) \sin(\omega t + \theta)]^2 = \text{minimum}. \quad (3)$$

Obliczamy pochodne cząstkowe funkcji (3):

$$\begin{aligned}
 \frac{\partial S}{\partial A} &= -2 \cdot \sum_{t=1}^n [z_t - (A + Bt) \sin(\omega t + \theta)] \cdot \sin(\omega t + \theta) , \\
 \frac{\partial S}{\partial B} &= -2 \cdot \sum_{t=1}^n [z_t - (A + Bt) \sin(\omega t + \theta)] \cdot t \sin(\omega t + \theta) , \\
 \frac{\partial S}{\partial \theta} &= -2 \cdot \sum_{t=1}^n [z_t - (A + Bt) \sin(\omega t + \theta)] \cdot (A + Bt) \cos(\omega t + \theta) , \\
 \frac{\partial S}{\partial \omega} &= -2 \cdot \sum_{t=1}^n [z_t - (A + Bt) \sin(\omega t + \theta)] \cdot (At + Bt^2) \cos(\omega t + \theta) .
 \end{aligned} \tag{4}$$

Korzystając z warunku koniecznego istnienia ekstremum funkcji F , po przekształceniach otrzymujemy układ równań normalnych dla modelu (2):

$$\left\{ \begin{aligned}
 \hat{A} \cdot \sum_{t=1}^n \sin^2(\hat{\omega} t + \hat{\theta}) + \hat{B} \cdot \sum_{t=1}^n t \sin^2(\hat{\omega} t + \hat{\theta}) &= \sum_{t=1}^n z_t \sin(\hat{\omega} t + \hat{\theta}) ; \\
 \hat{A} \cdot \sum_{t=1}^n t \sin^2(\hat{\omega} t + \hat{\theta}) + \hat{B} \cdot \sum_{t=1}^n t^2 \sin^2(\hat{\omega} t + \hat{\theta}) &= \sum_{t=1}^n t z_t \sin(\hat{\omega} t + \hat{\theta}) ; \\
 \frac{1}{2} \hat{A}^2 \cdot \sum_{t=1}^n \sin(2\hat{\omega} t + 2\hat{\theta}) + \hat{A} \hat{B} \cdot \sum_{t=1}^n t \sin(2\hat{\omega} t + 2\hat{\theta}) + \frac{1}{2} \hat{B}^2 \cdot \sum_{t=1}^n t^2 \sin(2\hat{\omega} t + 2\hat{\theta}) &= \\
 &= \hat{A} \cdot \sum_{t=1}^n z_t \cos(\hat{\omega} t + \hat{\theta}) + \hat{B} \cdot \sum_{t=1}^n t z_t \cos(\hat{\omega} t + \hat{\theta}) ; \\
 \frac{1}{2} \hat{A}^2 \cdot \sum_{t=1}^n t \sin(2\hat{\omega} t + 2\hat{\theta}) + \hat{A} \hat{B} \cdot \sum_{t=1}^n t^2 \sin(2\hat{\omega} t + 2\hat{\theta}) + \frac{1}{2} \hat{B}^2 \cdot \sum_{t=1}^n t^3 \sin(2\hat{\omega} t + 2\hat{\theta}) &= \\
 &= \hat{A} \cdot \sum_{t=1}^n t z_t \cos(\hat{\omega} t + \hat{\theta}) + \hat{B} \cdot \sum_{t=1}^n t^2 z_t \cos(\hat{\omega} t + \hat{\theta}) .
 \end{aligned} \right. \tag{5}$$

Układ równań (5) ma skomplikowaną postać. Rozwiązanie tego układu, czyli obliczenie wartości ocen parametrów A , B , ω , θ , jest zadaniem bardzo trudnym. Przewyciężenie tej trudności pozwoliłoby na wyznaczenie okresu wahań T na podstawie danych empirycznych. Jest to odrębne i niełatwe zagadnienie. W dalszym ciągu niniejszych rozważań nałożymy pewne ograniczenia na dane empiryczne. Ograniczenia te pozwolą znacznie uprościć układ równań (5) i dokonać estymacji szukanych parametrów. Będziemy od tego momentu za-

kładać, że znany jest okres \hat{T} (czyli także częstość $\hat{\omega}$) badanego zjawiska. Ma to miejsce przy opisie sezonowości, wtedy dla danych miesięcznych $\hat{T} = 12$, a dla kwartalnych $\hat{T} = 4$. Przyjmujemy także dwa następujące założenia: 1° dysponujemy kolejnymi punktami empirycznymi (z, t) ($t = 1, 2, \dots, n$), znaczy to, że szereg czasowy jest kompletny, nie ma brakujących danych empirycznych; 2° liczba punktów empirycznych jest całkowitą wielokrotnością okresu badanego zjawiska, tzn. $n = k\hat{T}$ – łatwo to zapewnić przez pominięcie w obliczeniach co najwyżej kilku punktów empirycznych.

Przy obecnej sprawozdawczości i szerokim dostępie do danych można uznać, że powyższe restrykcje nie powodują istotnych ograniczeń zastosowania proponowanych w tym artykule rozwiązań.

Dalsze rozważania będą miały na celu uproszczenie układu równań (5). Ze względu na to, że założyliśmy znajomość okresu \hat{T} , czwarte równanie w układzie (5) należy pominąć. Sumy występujące po lewych stronach pozostałych równań obliczamy wykorzystując odpowiednie wzory indukcyjne. W czterech przypadkach zostało to już pokazane w pracy [7], przytoczymy tutaj ostateczne wzory:

$$S_1 = \sum_{t=1}^n \sin(2\hat{\omega}t + 2\hat{\theta}) = \frac{\sin(n\hat{\omega}) \cdot \sin[(n+1)\hat{\omega} + 2\hat{\theta}]}{\sin \hat{\omega}},$$

$$S_2 = \sum_{t=1}^n \sin^2(\hat{\omega}t + \hat{\theta}) = \frac{n}{2} - \frac{\sin(n\hat{\omega}) \cdot \cos[(n+1)\hat{\omega} + 2\hat{\theta}]}{2\sin \hat{\omega}},$$

$$S_3 = \sum_{t=1}^n t \cdot \sin(2\hat{\omega}t + 2\hat{\theta}) = \frac{(n+1)\sin(2n\hat{\omega} + 2\hat{\theta}) - n\sin[2(n+1)\hat{\omega} + 2\hat{\theta}] - \sin 2\hat{\theta}}{4\sin^2 \hat{\omega}},$$

$$S_4 = \sum_{t=1}^n t \cdot \sin^2(\hat{\omega}t + \hat{\theta}) = \frac{n(n+1)}{4} - \frac{(n+1)\cos(2n\hat{\omega} + 2\hat{\theta}) - n\cos[2(n+1)\hat{\omega} + 2\hat{\theta}] - \cos 2\hat{\theta}}{8\sin^2 \hat{\omega}}.$$

Obliczenie kolejnych sum wymaga zastosowania wzoru, którego prawdziwości potrafimy dowieść. Wzór ten ma postać:

$$\sum_{k=1}^n k^2 a^k = \frac{n^2 a^{n+3} - (2n^2 + 2n - 1)a^{n+2} + (n^2 + 2n + 1)a^{n+1} - a^2 - a}{(a-1)^3}, \quad \text{dla } n \geq 1 \text{ i } a \neq 1.$$

Po podstawieniu do niego $a = \cos x + i \sin x$ i dokonaniu odpowiednich przekształceń, otrzymujemy dwa nowe i bardzo praktyczne wzory indukcyjne:

$$\sum_{k=1}^n k^2 \cos kx = \frac{-n^2 \sin\left(\frac{2n+3}{2}x\right) + (2n^2 + 2n - 1) \sin\left(\frac{2n+1}{2}x\right) - (n^2 + 2n + 1) \sin\left(\frac{2n-1}{2}x\right)}{8 \sin^3 \frac{x}{2}},$$

$$\sum_{k=1}^n k^2 \sin kx = \frac{n^2 \cos\left(\frac{2n+3}{2}x\right) - (2n^2 + 2n - 1) \cos\left(\frac{2n+1}{2}x\right) + (n^2 + 2n + 1) \cos\left(\frac{2n-1}{2}x\right) - 2 \cos \frac{x}{2}}{8 \sin^3 \frac{x}{2}}.$$

Wykorzystując te związki obliczamy dwie ostatnie sumy brakujące w układzie równań (5). Mamy więc:

$$S_5 = \sum_{t=1}^n t^2 \sin(2\hat{\omega}t + 2\hat{\theta}) = \\ = \frac{n^2 \cos[(2n+3)\hat{\omega} + 2\hat{\theta}] - (2n^2 + 2n - 1) \cos[(2n+1)\hat{\omega} + 2\hat{\theta}] + (n^2 + 2n + 1) \cos[(2n-1)\hat{\omega} + 2\hat{\theta}] - 2 \cos \hat{\omega} \cos 2\hat{\theta}}{8 \sin^3 \hat{\omega}},$$

$$S_6 = \sum_{t=1}^n t^2 \sin^2(\hat{\omega}t + \hat{\theta}) = \frac{n(n+1)(2n+1)}{12} + \\ + \frac{n^2 \sin[(2n+3)\hat{\omega} + 2\hat{\theta}] - (2n^2 + 2n - 1) \sin[(2n+1)\hat{\omega} + 2\hat{\theta}] + (n^2 + 2n + 1) \sin[(2n-1)\hat{\omega} + 2\hat{\theta}] - 2 \cos \hat{\omega} \sin 2\hat{\theta}}{16 \sin^3 \hat{\omega}}.$$

Jeśli uwzględnimy poczynione wcześniej założenie $n = k\hat{T}$, powyższe wzory przybiorą znacznie prostszą postać. Ponieważ $\hat{\omega} = 2\pi / \hat{T}$, więc $n\hat{\omega} = 2k\pi$, wobec tego po dokonaniu elementarnych przekształceń otrzymujemy:

$$S_1 = 0, \quad S_2 = \frac{n}{2}, \quad S_3 = \frac{-n \cos(\hat{\omega} + 2\hat{\theta})}{2 \sin \hat{\omega}}, \quad S_4 = \frac{n(n+1)}{4} - \frac{n \sin(\hat{\omega} + 2\hat{\theta})}{4 \sin \hat{\omega}},$$

$$S_5 = \frac{(n^2 + n) \sin 2\hat{\theta} - n^2 \cos \hat{\omega} \sin(\hat{\omega} + 2\hat{\theta})}{2 \sin^2 \hat{\omega}}, \quad S_6 = \frac{n(n+1)(2n+1)}{12} + \frac{n^2 \cos(2\hat{\omega} + 2\hat{\theta}) - (n^2 + 2n) \cos 2\hat{\theta}}{8 \sin^2 \hat{\omega}}.$$

Uwzględniając powyższe związki i pomijając czwarte równanie, układ równań normalnych zapisujemy w postaci:

$$\left\{ \begin{aligned} \hat{A} \cdot \frac{n}{2} + \hat{B} \cdot \left[\frac{n(n+1)}{4} - \frac{n \sin(\hat{\omega} + 2\hat{\theta})}{4 \sin \hat{\omega}} \right] &= \sum_{t=1}^n z_t \sin(\hat{\omega} t + \hat{\theta}); \\ \hat{A} \cdot \left[\frac{n(n+1)}{4} - \frac{n \sin(\hat{\omega} + 2\hat{\theta})}{4 \sin \hat{\omega}} \right] + \hat{B} \cdot \left[\frac{n(n+1)(2n+1)}{12} + \frac{n^2 \cos(2\hat{\omega} + 2\hat{\theta}) - (n^2 + 2n) \cos 2\hat{\theta}}{8 \sin^2 \hat{\omega}} \right] &= \sum_{t=1}^n t z_t \sin(\hat{\omega} t + \hat{\theta}); \\ \hat{A} \hat{B} \cdot \frac{-n \cos(\hat{\omega} + 2\hat{\theta})}{2 \sin \hat{\omega}} + \hat{B}^2 \cdot \frac{(n^2 + n) \sin 2\hat{\theta} - n^2 \cos \hat{\omega} \sin(\hat{\omega} + 2\hat{\theta})}{4 \sin^2 \hat{\omega}} &= \hat{A} \cdot \sum_{t=1}^n z_t \cos(\hat{\omega} t + \hat{\theta}) + \hat{B} \cdot \sum_{t=1}^n t z_t \cos(\hat{\omega} t + \hat{\theta}). \end{aligned} \right. \quad (6)$$

W układzie (6) ω jest znanym parametrem, a niewiadomymi są: A , B , θ . Układ ten, mimo swojej skomplikowanej postaci, daje możliwość estymacji parametrów A , B , θ . Można go rozwiązać wykorzystując odpowiedni program komputerowy, pozwalający na numeryczne rozwiązywanie układów równań.

Estymacja składnika wahań sezonowych

Układ równań (6) jest szczególnie przydatny w przypadku opisu zmienności sezonowej, ponieważ znamy wtedy okres wahań \hat{T} . W przypadku obserwacji miesięcznych $\hat{T} = 12$, czyli $\hat{\omega} = 2\pi / \hat{T} = \pi / 6$, a w przypadku obserwacji kwartalnych $\hat{T} = 4$, więc $\hat{\omega} = 2\pi / \hat{T} = \pi / 2$. Gdy dysponujemy obserwacjami miesięcznymi, po podstawieniu $\hat{\omega} = \pi / 6$ i niezbędnych uproszczeniach, układ równań (6) przybiera postać:

$$\left\{ \begin{aligned} \hat{A} \cdot \frac{n}{2} + \hat{B} \cdot \left[\frac{n(n+1)}{4} - \frac{n}{4} \cos 2\hat{\theta} - \frac{n\sqrt{3}}{4} \sin 2\hat{\theta} \right] &= \sum_{t=1}^n z_t \sin\left(\frac{\pi}{6} t + \hat{\theta}\right); \\ \hat{A} \cdot \left[\frac{n(n+1)}{4} - \frac{n}{4} \cos 2\hat{\theta} - \frac{n\sqrt{3}}{4} \sin 2\hat{\theta} \right] + \hat{B} \cdot \left[\frac{n(n+1)(2n+1)}{12} - \frac{n(n+4)}{4} \cos 2\hat{\theta} - \frac{n^2\sqrt{3}}{4} \sin 2\hat{\theta} \right] &= \\ &= \sum_{t=1}^n t z_t \sin\left(\frac{\pi}{6} t + \hat{\theta}\right); \\ \hat{A} \hat{B} \cdot \left[\frac{n}{2} \sin 2\hat{\theta} - \frac{n\sqrt{3}}{2} \cos 2\hat{\theta} \right] + \hat{B}^2 \cdot \left[\frac{n(n+4)}{4} \sin 2\hat{\theta} - \frac{n^2\sqrt{3}}{4} \cos 2\hat{\theta} \right] &= \\ &= \hat{A} \cdot \sum_{t=1}^n z_t \cos\left(\frac{\pi}{6} t + \hat{\theta}\right) + \hat{B} \cdot \sum_{t=1}^n t z_t \cos\left(\frac{\pi}{6} t + \hat{\theta}\right). \end{aligned} \right. \quad (7)$$

Rozważany i oszacowany model (2) z kolei wykorzystamy do opisu i prognozy skupu mleka w Polsce. Jednocześnie, wykorzystując te same dane empiryczne, oszacujemy parametry modelu o stałej amplitudzie:

$$z_t = s + A \cdot \sin(\omega t + \theta) + \varepsilon_t. \quad (8)$$

Estymację parametrów modelu (8) przeprowadzono w pracy [7]. Gdy okres \hat{T} opisywanego zjawiska jest znany i $n = k\hat{T}$, wtedy oszacowanie parametrów modelu (8) jest bardzo proste. Mamy mianowicie:

$$\begin{aligned} \hat{s} &= \frac{1}{n} \cdot \sum_{t=1}^n z_t; & \hat{\theta} &= \arctg \left[\frac{\sum_{t=1}^n z_t \cos \hat{\omega} t}{\sum_{t=1}^n z_t \sin \hat{\omega} t} \right]; \\ \hat{A} &= \frac{2}{n} \cdot \left[\cos \hat{\theta} \cdot \sum_{t=1}^n z_t \sin \hat{\omega} t + \sin \hat{\theta} \cdot \sum_{t=1}^n z_t \cos \hat{\omega} t \right]. \end{aligned} \quad (9)$$

Zastosowanie praktyczne proponowanego modelu

W tabeli 1 podano dane dotyczące skupu mleka w Polsce w latach 1969–1979. Są to dane miesięczne (w mln litrów), które zaczerpnięto z pracy [7, s. 455] oraz z Biuletynów Statystycznych z różnych lat. Korzystamy z tych danych, gdyż wahania sezonowe skupu mleka w latach 1969–1979 charakteryzowały się właśnie rosnącą amplitudą. Ponadto, chcemy dokonać porównania z wynikami otrzymanymi w pracy [7].

Do estymacji parametrów modelu (2) wykorzystujemy dane dotyczące skupu mleka w Polsce z lat 1970–1977, pozostałe dane posłużą do oceny wyznaczonych prognoz. Trend opisujemy funkcją liniową $f(t) = a + bt$. Jej parametry szacujemy metodą średnich. Odpowiedni układ równań ma postać:

$$1176 \cdot \hat{b} + 48 \cdot \hat{a} = 24734,9,$$

$$3480 \cdot \hat{b} + 48 \cdot \hat{a} = 34680,1.$$

Po jego rozwiązaniu i zaokrągleniach otrzymujemy:

$$\hat{f}(t) = 409,5563 + 4,3165 \cdot t, \quad \text{gdzie } t = 1 \equiv \text{styczeń 1970 r.}$$

Następnie obliczamy różnice $z_t = y_t - \hat{f}(t)$ dla $t = 1 \equiv \text{I 1970 r.}$ do $t = 96 \equiv \text{XII 1977 r.}$ i za ich pomocą obliczamy potrzebne sumy:

Tabela 1
Skup mleka w Polsce w latach 1969–1979 (w mln l)

Miesiąc \ Rok	Rok										
	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
I	346,8	336,5	353,1	410,7	476,6	564,6	514,9	588,9	641,5	656,8	618,0
II	328,1	319,6	339,1	407,8	450,4	527,5	481,3	561,8	604,4	605,3	557,8
III	376,9	369,8	381,4	457,5	522,1	596,8	555,5	633,0	704,7	713,6	680,4
IV	380,8	391,3	392,7	473,9	520,3	594,1	581,2	635,1	715,9	741,6	713,0
V	470,6	475,1	522,4	637,6	699,6	752,3	783,3	806,4	933,4	953,2	895,2
VI	590,1	596,4	612,2	726,1	806,5	911,1	907,3	983,8	1043,7	1113,0	1055,0
VII	573,4	591,5	597,8	682,6	777,1	888,9	879,6	946,7	1073,5	1123,0	1122,7
VIII	481	559,4	525,0	654,6	744,0	807,9	828,3	921,1	994,5	1043,9	1073,1
IX	444,2	523,4	507,2	638,7	710,4	759,1	784,4	827,6	887,1	922,3	980,3
X	419,5	463,5	470,9	599,5	638,9	660,7	702,8	720,5	818,2	814,9	912,0
XI	326	351,1	375,1	469,2	514,3	515,2	541,0	612,7	642,1	674,8	691,1
XII	303,4	329,4	379,4	439,2	514,0	484,8	550,1	594,4	616,4	637,6	679,2
Skup roczny	5040,8	5307,0	5456,3	6597,4	7374,2	8063,0	8109,7	8832,0	9675,4	10000,0	9977,8

Źródło: [7, s. 455] i opracowanie własne na podstawie Biuletynów Statystycznych z różnych lat.

$$\sum_{t=1}^{96} z_t \sin\left(\frac{\pi}{6}t\right) = -3820,8837; \quad \sum_{t=1}^{96} z_t \cos\left(\frac{\pi}{6}t\right) = -7258,3966;$$

$$\sum_{t=1}^{96} t z_t \sin\left(\frac{\pi}{6}t\right) = -194246,6642; \quad \sum_{t=1}^{96} t z_t \cos\left(\frac{\pi}{6}t\right) = -395983,5646.$$

Po podstawieniu ich do (7) układ równań normalnych w naszym przykładzie przybiera następującą postać:

$$\begin{cases} 48 \cdot \hat{A} + (2328 - 24 \cdot \cos 2\hat{\theta} - 41,5692 \cdot \sin 2\hat{\theta}) \cdot \hat{B} = -3820,8837 \cdot \cos \hat{\theta} - 7258,3966 \cdot \sin \hat{\theta}; & (10) \\ (2328 - 24 \cdot \cos 2\hat{\theta} - 41,5692 \cdot \sin 2\hat{\theta}) \cdot \hat{A} + \\ + (149768 - 2400 \cdot \cos 2\hat{\theta} - 3990,6451 \cdot \sin 2\hat{\theta}) \cdot \hat{B} = -194246,6642 \cdot \cos \hat{\theta} - 395983,5646 \cdot \sin \hat{\theta}; \\ (48 \cdot \sin 2\hat{\theta} - 83,1384 \cdot \cos 2\hat{\theta}) \cdot \hat{A} \cdot \hat{B} + (2400 \sin 2\hat{\theta} - 3990,6451 \cdot \cos 2\hat{\theta}) \cdot \hat{B}^2 = \\ = -7258,3966 \cdot \hat{A} \cdot \cos \hat{\theta} + 3820,8837 \cdot \hat{A} \cdot \sin \hat{\theta} - 395983,5646 \cdot \hat{B} \cdot \cos \hat{\theta} + 194246,6642 \cdot \hat{B} \cdot \sin \hat{\theta}. \end{cases}$$

Poszukujemy jego rozwiązania dla $\hat{\theta} \in [-\pi; \pi]$.

Układ równań (10) rozwiązano przy użyciu programu komputerowego *Mathematica 3.0*. Z rozwiązań równoważnych wybrano:

$$\hat{A} = 110,24, \quad \hat{B} = 1,26178, \quad \hat{\theta} = -2,05727.$$

Ostatecznie, po zaokrągleniach, oszacowany model dla skupu mleka w Polsce przyjmuje postać:

$$\hat{y}(t) = 409,57 + 4,317 \cdot t + (110,24 + 1,262 \cdot t) \cdot \sin\left(\frac{\pi}{6}t - 2,057\right), \quad (11)$$

w którym $t = 1 \equiv$ styczeń 1970 r.

Miary dopasowania modelu (11) w przedziale jego aproksymacji kształtują się następująco:

$$\text{współczynnik determinacji } R^2 = 1 - \left[\frac{\sum_{t=1}^{96} (y_t - \hat{y}_t)}{\sum_{t=1}^{96} (y_t - \bar{y})} \right] = 0,9404,$$

który interpretujemy w ten sposób, że oszacowany model (11) objaśnia w 94% zmienność badanej zmiennej zależnej; odchylenie standardowe składnika

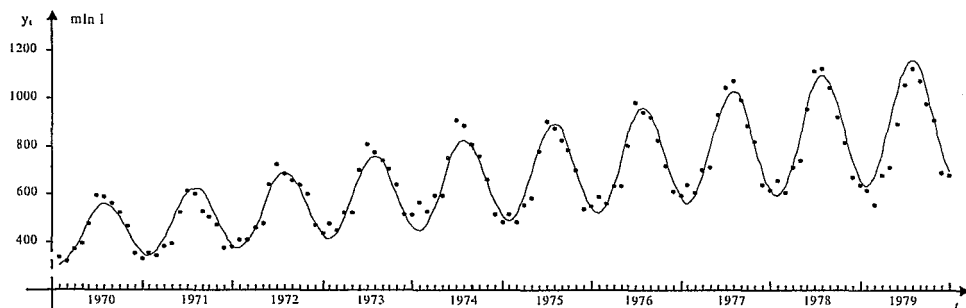
resztowego $s_e = \sqrt{\sum_{t=1}^{96} (y_t - \hat{y}_t)^2 / (96-5)} = 44,8$ mln l; współczynnik zmienności przypadkowej $v = 100s_e/\bar{y} = 7,3\%$.

Na rysunku 1 zamieszczono dane empiryczne i wykres oszacowanego modelu (11).

Dla porównania oszacujemy jeszcze model (8) o stałej amplitudzie. Funkcja trendu jest taka sama jak poprzednio, a parametry składnika wahań sezonowych szacujemy za pomocą wzorów (9). Otrzymujemy: $\hat{s} = 16 \cdot 10^{-9}$, $\hat{\theta} = 1,086246$, $\hat{A} = -170,88855$. Przekonujemy się więc, że estymacja parametrów trendu metodą średnich powoduje zmniejszenie o jeden liczby parametrów harmoniki. Wtedy we wzorach (9) zawsze otrzymujemy $\hat{s} = 0$. Ostatecznie, po zaokrągleniach, model ze stałą amplitudą przyjmuje postać:

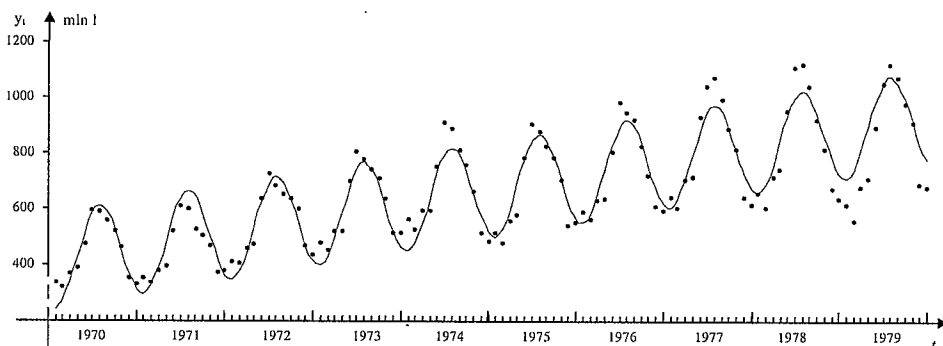
$$\hat{y}(t) = 409,56 + 4,317 \cdot t + 170,889 \cdot \sin\left(\frac{\pi}{6}t - 2,056\right), \quad (12)$$

w którym $t = 1 \equiv$ styczeń 1970 r. Miary dopasowania modelu (12) do tych samych danych empirycznych są następujące: $R^2 = 0,921$, $s_e = 51,5$, $v = 8,3\%$, a jego wykres zamieszczono na rysunku 2.



Rysunek 1

Skup mleka w Polsce w latach 1970–1979 (model o rosnącej amplitudzie, prognoza na lata 1978 i 1979)



Rysunek 2

Skup mleka w Polsce w latach 1970–1979 (model o stałej amplitudzie, prognoza na lata 1978 i 1979)

Prognozy obliczone na podstawie wzorów (11) i (12) na rok 1978 oraz kres górny błędu względnego wyrażony w procentach dla każdego miesiąca podane są w tabelach 2 oraz 3. W tabeli 4 pokazana została także prognoza wsteczna na rok 1969 obliczona na podstawie (11).

Tabela 2

Prognoza na rok 1978 otrzymana za pomocą modelu o rosnącej amplitudzie (11)

1978	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
y_t	656,8	605,3	713,6	741,6	953,2	1113,0	1123,0	1043,9	922,3	814,9	674,8	637,6
\hat{y}_t	595,8	634,5	727,0	850,0	971,9	1061,1	1094,2	1063,0	976,3	858,0	741,0	657,8
Kgbw%	9,29	4,82	1,87	14,61	1,96	4,67	2,57	1,83	5,85	5,29	9,81	3,17

Tabela 3

Prognoza na rok 1978 otrzymana za pomocą modelu o stałej amplitudzie (12)

1978	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
y_t	656,8	605,3	713,6	741,6	953,2	1113,0	1123,0	1043,9	922,3	814,9	674,8	637,6
\hat{y}_t	657,5	688,0	757,3	847,9	936,7	1001,0	1024,9	1003,0	942,4	860,4	780,3	724,5
Kgbw%	0,11	13,67	6,12	14,33	1,73	10,06	8,73	3,92	2,18	5,59	15,63	13,63

Tabela 4

Prognoza wsteczna na rok 1969 otrzymana za pomocą modelu o rosnącej amplitudzie (11)

1969	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
y_t	346,8	328,1	376,9	380,8	470,6	590,1	573,4	481,0	444,2	419,5	326,0	303,4
\hat{y}_t	265,8	283,7	324,5	378,7	433,3	474,4	491,8	481,4	446,4	396,9	347,3	312,1
Kgbw%	23,36	13,53	13,91	0,54	7,93	19,60	14,23	0,08	0,49	5,38	6,53	2,87

Podsumowanie

Model (11) o rosnącej monotonicznie amplitudzie sezonowości jest odpowiednim modelem do opisu skupu mleka w Polsce, odzwierciedla zasadnicze cechy tej zmiennej zależnej, ale nie eliminuje oczywistych w tej sytuacji wahań przypadkowych. Stanowi pewien postęp w stosunku do przytoczonego modelu (12) o stałej amplitudzie sezonowości. Zapewnia zadowalającą prognozę długoterminową. Model (12) ma bardzo prostą estymację parametrów (9) i był wielokrotnie wykorzystywany do opisu odpowiadających mu teoretycznie zjawisk klimatycznych. Estymacja parametrów modelu (2) ze zmienną amplitudą jest zadaniem złożonym i pracochłonnym, ale wyprowadzenie odpowiadających mu równań normalnych (6) i ich uproszczeń przyszłego użytkownika nie musi interesować, może skorzystać z ostatecznych propozycji (7). Pamiętajmy też o tym, że ten sam problem można rozwiązywać w różny sposób, ale zawsze należy stosować metody zgodne z surowymi regułami postępowania statystycznego.

Literatura

- BOX G.E.P., JENKINS G.M., 1983: *Analiza szeregów czasowych. Prognoza i sterowanie*. PWN, Warszawa. [1]
 CHOW G.C., 1995: *Ekonometria*. PWN, Warszawa. [2]
 CIEŚLAK M. (red. naukowa), 1997: *Prognozowanie gospodarcze. Metody i zastosowania*. PWN, Warszawa. [3]
 MIŁO W., 1990: *Nieliniowe modele ekonometryczne*. PWN, Warszawa. [4]
 RADZIKOWSKA B. (red.), 2000: *Metody prognozowania. Zbiór zadań*. Wydawnictwo AE we Wrocławiu, Wrocław. [5]

- RYŻYK I.M., GRADSZTEJN I.S., 1964: *Tablice całek, sum, szeregów i iloczynów*. PWN, Warszawa. [6]
- SMOLIK S., 1995: *Uproszczona procedura estymacji modelu wahań okresowych*. „Przegląd Statystyczny”, R. XLII, z. 3–4. [7]
- SMOLIK S., 1998: *Oszczędne modele dla okresowych szeregów czasowych*. [w:] *Przestrzenno-czasowe modelowanie i prognozowanie zjawisk gospodarczych*. Red. A. Zeliaś, Akademia Ekonomiczna w Krakowie, s.177–188. [8]
- WELFE W., WELFE A., 1996: *Ekonometria stosowana*. PWN, Warszawa. [9]
- ZAJĄC K., 1994: *Zarys metod statystycznych*, PWE, Warszawa. [10]
- ZELIAŚ A., 1997: *Teoria prognozy*. Wyd. III zmienione, PWE, Warszawa. [11]

Description of periodic phenomena with a monotonically changing amplitude

Abstract

For the purposes of description of economic and natural phenomena with a distinctly outlined trend and a dominating seasonal dependence, the model was proposed: $y_t = f(t) + (A + B \cdot t) \cdot \sin(\omega t + \theta) - \varepsilon_t$ ($t = 1, 2, \dots, n$). The analytic shape of the trend function $f(t)$, in its general form, is adjusted to the phenomenon being described and subsequently, its parameters are estimated with the average method, using empirical data (t, y_t) . In such a case, seasonal relations with changing amplitude can sufficiently be described with the model:

$$z_t = (A + B \cdot t) \cdot \sin(\omega t + \theta) - \varepsilon_t, \text{ whereas } z_t = y_t - \hat{f}(t) \quad (t = 1, 2, \dots, n)$$

and $\omega = 2\pi/T$. Whilst using monthly observations $\hat{T} = 12$ and $n = k\hat{T}$, this model's parameters are liable to estimation out of the following equation system:

$$\left\{ \begin{array}{l}
 \hat{A} \cdot \frac{n}{2} + \hat{B} \cdot \left[\frac{n(n+1)}{4} - \frac{n}{4} \cos 2\hat{\theta} - \frac{n\sqrt{3}}{4} \sin 2\hat{\theta} \right] = \sum_{t=1}^n z_t \sin\left(\frac{\pi}{6}t + \hat{\theta}\right); \\
 \hat{A} \cdot \left[\frac{n(n+1)}{4} - \frac{n}{4} \cos 2\hat{\theta} - \frac{n\sqrt{3}}{4} \sin 2\hat{\theta} \right] + \hat{B} \cdot \left[\frac{n(n+1)(2n+1)}{12} - \frac{n(n+4)}{4} \cos 2\hat{\theta} - \frac{n^2\sqrt{3}}{4} \sin 2\hat{\theta} \right] = \\
 \hspace{25em} = \sum_{t=1}^n t z_t \sin\left(\frac{\pi}{6}t + \hat{\theta}\right); \\
 \hat{A}\hat{B} \cdot \left[\frac{n}{2} \sin 2\hat{\theta} - \frac{n\sqrt{3}}{2} \cos 2\hat{\theta} \right] + \hat{B}^2 \cdot \left[\frac{n(n+4)}{4} \sin 2\hat{\theta} - \frac{n^2\sqrt{3}}{4} \cos 2\hat{\theta} \right] = \\
 \hspace{15em} = \hat{A} \cdot \sum_{t=1}^n z_t \cos\left(\frac{\pi}{6}t + \hat{\theta}\right) + \hat{B} \cdot \sum_{t=1}^n t z_t \cos\left(\frac{\pi}{6}t + \hat{\theta}\right).
 \end{array} \right.$$

This model has been employed, with a linear tendency, in our description and forecast of milk purchase in Poland.