Metoda $\bar{r}$
– substytutem metody najmniejszych kwadratów

Igor Timofiejuk, prof. zw. dr hab.
Katedra Ekonomii Politycznej WNE UW

W szeregu chronologicznym (czasowym, dynamicznym, rozwojowym)[1] można wyróżnić trzy elementy składowe[2]:

1) tendencja rozwojowa (trend),
2) wahania regularne (okresowe),
3) wahania nieregularne (przypadkowe).

Dwa ostatnie składniki nie będą tu przedmiotem rozważań, a poświęcimy je problemowi wyrażania tendencji rozwojowej (trendu). Tendencja rozwojowa (trend) to właściwie istota procesu lub zjawiska, a więc prawidłowość ewolucyjna powodująca wzrost, stagnację czy też spadek wartości liczbowych procesu, odzwierciedlanych w wyniku pomiaru, w szeregu chronologicznym.

Tendencję rozwojową możemy określić na trzy sposoby:

1) licząc średnie ruchome przy określeniu liczby wyrazów szeregu czasowego, w których powtarzają się wahania regularne. Metodę tę określa się mianem mechanicznej;
2) analizując zachowanie się wyrazów szeregu chronologicznego i w wyniku tego obierając rodzaj funkcji matematycznej odzwierciedlającej zachowanie wyrazów szeregu i w efekcie dopasowanie tej funkcji metodą najmniejszych kwadratów do danych owego szeregu. Tę metodę określa się mianem analitycznej;
3) licząc przeciętne (średnie) tempo (stopę) wzrostu, np. metodą średniej geometrycznej ważonej systemem wag jednostkowych (krótko: nieważonej średniej geometrycznej — r_g) lub mojej metody, uwzględniającej sumę

Artykuł jednocześnie ukazuje się w 2/2204 „Ekonomiczno-Informatycznego Kwartalnika Teoretycznego" Wyższej Szkoły Ekonomiczno-Informatycznej.

[1] Ujmując w kategoriach związku funkcyjnego lub korelacyjnego można wyrazić to tak: jeśli wielkościom czasu (momentom lub okresom), tzn. czynnikowi egzogenicznemu względem badanego procesu, przypiszemy adekwatnie wielkości liczbowe charakteryzujące ów proces, to wówczas otrzymamy szereg chronologiczny (czasowy, rozwojowy, dynamiczny) momentów lub okresów. Jak wiadomo, szereg okresów przedstawia s t r u m i e n i e, a szereg momentów — z a s o b y.

[2] Inną sprawą jest ujmowanie związku tych elementów składowych, tzn. czy zachodzą tu zależności addytywne, czy też multiplikatywne.

wyrazów szeregu czasowego ($\bar{r}$)[3]. Ten sposób wyrażania trendu też trzeba określić mianem analityczny.

Ogólnie, wszystkie te sposoby wyrażania trendu określa się jako wyrównywanie (wygładzanie) szeregów czasowych.

I

Mechaniczna metoda średniej ruchomej (lub scentrowanej średniej ruchomej, gdy liczba wyrazów mierzących wahania regularne jest parzysta) eliminuje wahania okresowe. Ale poza trendem pozostają w danych wygładzanych także wahania przypadkowe. Czyni to trend mało eleganckim, nieco „chropowatym”. Owe chropowatości powstają w wyniku wahań nieregularnych i mogą być dobrym impulsem ku zwróceniu uwagi badacza na elementy (składniki) procesu niepoznane lub poznane nie do końca. W tym sensie średnia ruchoma może spełniać się jako impuls inspirujący.

II

Analityczna metoda najmniejszych kwadratów, stworzona — niezależnie — przez C.F. Gaussa (1777–1856) i A.M. Legendre'a (1752–1833), to metoda wyrównywania danych doświadczalnych, mająca właściwość minimalizowania wielkości błędów pomiaru.

Jeśli mamy szereg chronologiczny:

$t_0, t_1, t_2, \ldots, t_{n-1}, t_n$ — czas (zmienna niezależna)
$x_0, x_1, x_2, \ldots, x_{n-1}, x_n$ — zmienna zależna

to jego wartości liczbowe $x_0, x_1, x_2, \ldots, x_{n-1}, x_n$ określamy jako dane r z e c z y w i s t e (e k s p e r y m e n t a l n e). Po analizie danych rzeczywistych wybieramy funkcję klasy elementarnej, np. liniową, potęgową, wykładniczą, logarytmiczną czy też funkcje trygonometryczne[4]; w tu badanym zagadnieniu $y = f(t)$, tzn. argumentem (zmienną niezależną) jest czas. Wartości liczbowe tej funkcji $y_0, y_1, y_2, \ldots, y_{n-1}, y_n$ noszą nazwę d a n y c h t e o r e t y c z n y c h (w y l i c z o n y c h) i dopasowujemy je, właśnie metodą najmniejszych kwadratów, do danych rzeczywistych. Istota metody najmniejszych kwadratów zawiera się w wyrażeniu:

$$\sum_{i=0}^{n} (x_i - y_i)^2 = \min$$

[3] Na temat metody $\bar{r}$ opublikowałem dwie książki i szereg artykułów naukowych. Dla potrzeb tego opracowania wystarczające jest wskazanie na I. Timofiejuk, *Szeregi czasowe — pomiar przeciętnej dynamiki*, „Ekonomia” 2001 nr 2.

[4] Por. G.M. Fichtenholz, *Rachunek różniczkowy i całkowy*, t. I, wyd. czwarte fotograficzne, PWN, Warszawa 1966, s. 81–85.

tzn. takim, gdzie suma kwadratów odchyleń danych rzeczywistych od danych teoretycznych jest minimalna. Nie ma innej techniki matematycznej minimalizującej odchylenia danych eksperymentalnych od wyliczonych poza metodą najmniejszych kwadratów.

Metodę najmniejszych kwadratów cechują dwie właściwości:

a) nie wymaga żadnych założeń co do funkcji, którą dopasowujemy do danych rzeczywistych; wybór wynika z analizy danych;
b) nie wymaga żadnych założeń co do równości wyrazów rzeczywistych i teoretycznych, tzn. żadna z wartości liczbowych teoretycznych nie musi być identyczna z którąkolwiek z wartości liczbowych eksperymentalnych.

III

Przeciętne (średnie) tempo (stopa) wzrostu[5] jako sposób określania tendencji rozwojowej (trendu) charakteryzuje się pewnymi założeniami. Otóż, średnie (przeciętne) tempo (stopa) wzrostu wyraża, dla dowolnej liczby okresów lub momentów, jedną liczbą dynamikę (prędkość) zmian mierników ekonomicznych. Nie zależy to od metody jego rachunku, tzn. czy jest to metoda średniej geometrycznej ważonej systemem wag jednostkowych, zwanej krótko i powszechnie nieważoną średnią geometryczną (r_g), czy też — wykorzystującej sumę wyrazów badanego szeregu chronologicznego (moja metoda — $\bar{r}$). Takie postępowanie zasadza się na:

a) przyjęciu zasady progresji (postępu geometrycznego) odzwierciedlanego w funkcji wykładniczej typu $y_i = a \cdot b^{t_i}$,
b) równości wartości liczbowych wybranych wyrazów szeregu rzeczywistego i teoretycznego: w metodzie r_g — wyrazu podstawowego (bazowego) i końcowego badanego szeregu czasowego; w metodzie $\bar{r}$ — wyrazu podstawowego (bazowego) i sumy wyrazów badanego szeregu dynamicznego.

Sumując, metody rachunku średniego tempa wzrostu mają mniej stopni „swobody" odzwierciedlania kształtu tendencji rozwojowej. Musi to być zawsze kształt funkcji wykładniczej o różnej li tylko skali „stromizny" wzrostu. Rozpoczynamy od nieważonej średniej geometrycznej — r_g.

Jeśli mamy szereg chronologiczny:

$$t_0, t_1, t_2, \ldots, t_{n-1}, t_n$$
$$Y_0, Y_1, \ldots, Y_{n-1}, Y_n$$

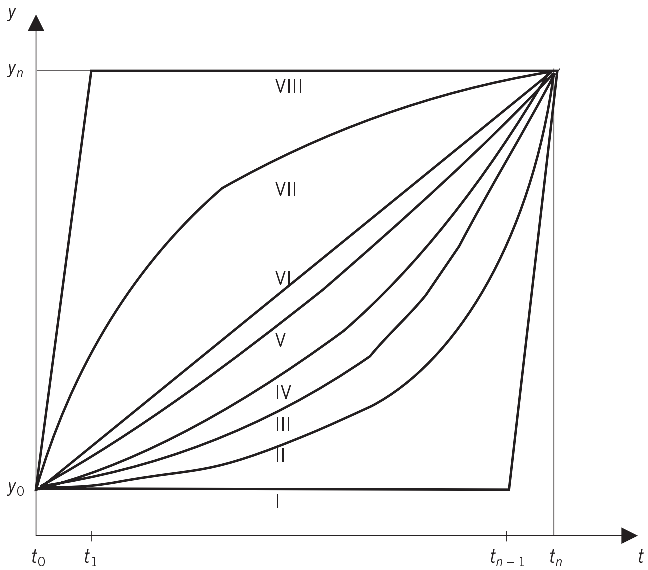
to średnie tempo wzrostu liczymy według formuły:

$$r_g = \sqrt[n]{Y_n : Y_0} - 1$$

---

[5] Przeciętne (średnie) tempo (stopa) wzrostu przyjęto powszechnie jako termin w ekonomii politycznej i *economics*. W istocie powinno się mówić, ściśle rzecz biorąc, o tempie (stopie) przy-

Wyliczone tym sposobem przeciętne tempo wzrostu ma charakter „nieokreślony" na podobieństwo teorii indeksów zespołowych typu Laspeyresa i Paaschego, a mianowicie jest to twierdzenie warunkowe, tzn. jaki musiałby być r ó w n o m i e r n y wzrost o ilorazie postępu geometrycznego $1 + r_g$, aby startując z poziomu Y_0 po n okresach lub momentach uzyskać wartość liczbową Y_n. Ale w rzeczywistości realne procesy nie rozwijają się „idealnie" równomiernie. Pod wpływem wielu czynników i tych z wahań regularnych, i tych ze zdarzeń przypadkowych (np. katastroficznych) realne procesy gospodarcze mogą przyspieszać w badanym odcinku czasu dynamikę ponad $1 + r_g$ lub ją zwalniać poniżej indeksu dynamiki $1 + r_g$. Tak więc rysuje się cała „wiązka" trajektorii, które startując z Y_0 dochodzą po „swojemu" do mety o wartości liczbowej Y_n. Niech o tym świadczy rysunek.



Rys. 1.

Hipotetyczne trajektorie startu z Y_0 do mety w Y_n według tego samego przeciętnego r_g

Źródło: koncepcja własna.

Trajektorie od I do VIII istotnie się różnią, ale ich przeciętne tempo wzrostu jest identyczne. Stąd wniosek zasadniczy: metoda r_g różni się od metody najmniejszych kwadratów tym, iż dane teoretyczne pozyskane z metody najmniejszych kwadratów mogą tyczyć się li tylko jednych jedynych danych rzeczywistych, natomiast dane teoretyczne z metody r_g według $Y_0(1 + r_g)^n$

rostu, tzn. gdy mamy x_i i x_{i+1} oraz $\Delta x_{i+1} = x_{i+1} - x_i$, czyli przyrost $\Delta x_{i+1} : x_i$ jest tempem przyrostu. Jednak przyjęło się to nazywać tempem wzrostu; nie ma w tym nic złego i kruszyć kopie o tę sprawę byłoby purystycznym kretynizmem. Tak często bywa w nauce, że niefortunne terminy przyjmują się i funkcjonują, np. termin funkcje regresji.

mogą odzwierciedlać w istocie nieskończoną „wiązkę” rzeczywistych trajektorii dynamiki. Zatem, metoda rachunku średniego tempa wzrostu wedle nieważonej średniej geometrycznej nie może być uważana, z wyżej wyłuszczonych powodów, za substytut metody najmniejszych kwadratów.

IV

Moja metoda, nazwana $\bar{r}$, a uwzględniająca sumę wyrazów badanego szeregu chronologicznego zasadza się na tych samych regułach progresji geometrycznej co i metoda r_g, z tym tylko, że wymaga, aby dynamika, startując z poziomu Y_0, w efekcie prowadziła do tej samej sumy wyrazów szeregu teoretycznego co suma wyrazów szeregu rzeczywistego. W tym sensie metoda $\bar{r}$ jest także twierdzeniem typu warunkowego typu: jaka musi być przeciętna stopa wzrostu, aby z poziomu wyjściowego (wyraz bazowy), idąc ku wyrazowi końcowemu $\bar{Y}_n$ (to wyraz teoretyczny), zachować sumę wyrazów szeregu teoretycznego identyczną z sumą wyrazów szeregu rzeczywistego, czyli:

$$\sum_{i=1}^{n} Y_i = a_n$$

Metoda $\bar{r}$ to rozwiązanie wielomianu lub funkcji całkowitej wymiernej o postaci:

$$f(x) = a_0 x^n + a_1 x^{n-1} + \ldots + a_{n-2} x^2 + a_{n-1} x + a_n$$

która, dla warunków postawionego zadania, sprowadza się do zredukowanej postaci:

$$f(\bar{x}) = \bar{x}^n + \bar{x}^{n-1} + \ldots + \bar{x}^2 + \bar{x} - a_n$$

Liczba pierwiastków tego równania, na mocy twierdzenia Descartesa-Harriota, dla konkretnych warunków tu rozważanych wynosi dokładnie jeden, w ciągu a_0, a_1, …, a_{n-1}, $-a_n$ jest bowiem tylko jedna zmiana znaku, a mianowicie przed a_n, który to wyraz jest ilorazem

$$\sum_{i=1}^{n} Y_i : Y_0$$

I to właśnie jest indeks (wskaźnik) wzrostu $\bar{x}$, a zatem na mocy równania $\bar{r} = \bar{x} - 1$ otrzymuje się przeciętną stopę wzrostu[6]. Praktycznie rachunek stopy wzrostu odbywa się na podstawie tablic[7].

Jak zatem można określić możliwość substytucji przez metodę $\bar{r}$ metody najmniejszych kwadratów?

---

[6] Zob. I. Timofiejuk, *Metoda $\bar{r}$. Teoria i tablice*, Fundacja Naukowa Taylora, Warszawa 1993, s. VII–X lub odpowiednio w języku angielskim XXIII–XXIV i rosyjskim XXXIX–XLII.

[7] Jak wyżej, dalej, powołując się na tę książkę, będę pisał *Tablice*.

V

Odpowiedź na wyżej postawione pytanie będzie udzielona w dwóch wariantach:

a) rzeczywiste Y_0 i $\sum_{i=1}^{n} Y_i$ i jakie ma być $\overline{Y}_n$?

b) rzeczywiste Y_n i $\sum_{i=1}^{n} Y_i$ i jakie ma być $\overline{Y}_0$?

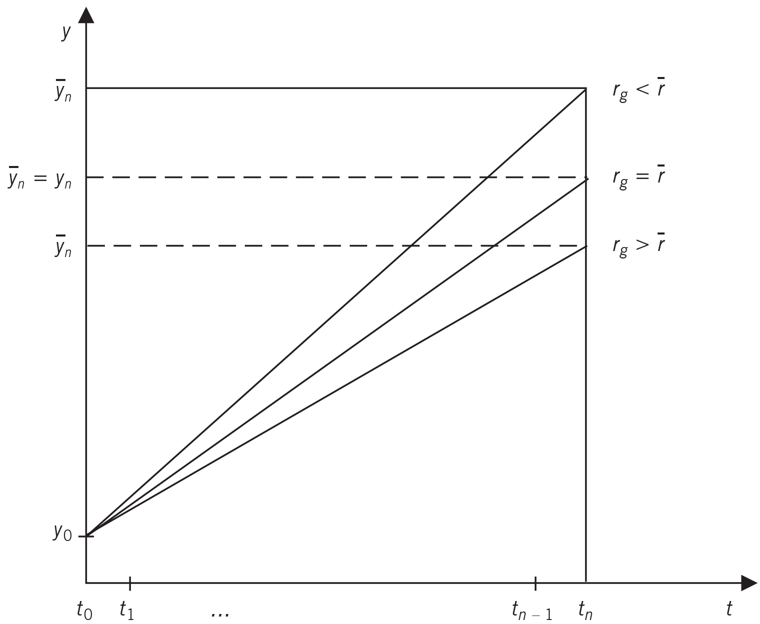
Przechodząc do wariantu pierwszego, tzn. poszukiwania $\overline{Y}_n$, przy zachowaniu rzeczywistych (eksperymentalnych) danych wyrazu początkowego (Y_0) i rzeczywistej sumy $\left(\sum_{i=1}^{n} Y_i\right)$ wyrazów szeregu czasowego, odpowiedź jest prosta. Jest tylko jedna trajektoria wyliczona według metody $\bar{r}$, która może przybrać trzy konfiguracje:

a) $\bar{r} = r_g$, wzrost równomierny,

b) $\bar{r} > r_g$, niedoszacowanie dynamiki przez metodę r_g,

c) $\bar{r} < r_g$, przeszacowanie dynamiki przez metodę r_g.

Poniższy rysunek wyjaśnia to nad wyraz wyraziście.



Rys. 2.

Relacje r_g i $\bar{r}$

Źródło: koncepcja własna.

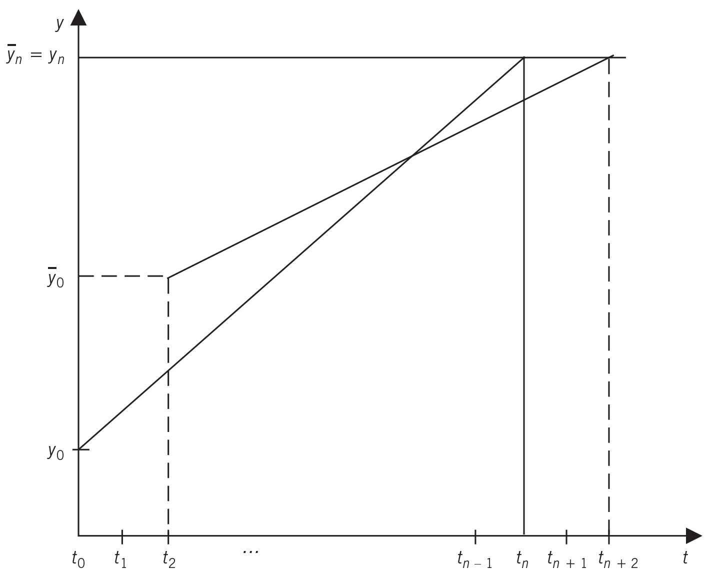
Tu sprawa jest jasna i jednoznaczna. Metoda $\bar{r}$ daje tylko jedną trajektorię teoretycznego wzrostu, zależną od tego, jaki jest układ relacji między $\bar{r}$ i r_g. Tu metoda $\bar{r}$ jest właściwie substytutem metody najmniejszych kwadratów w obrębie funkcji wykładniczej, która leży u podstaw rachunku przeciętnego (średniego) tempa (stopy) wzrostu.

VI

Przy wariancie $\bar{Y}_n$ i $\sum_{i=1}^{n} Y_i$ jako zachowane z danych rzeczywistych i poszukiwania $\bar{Y}_0$, które musi być, aby ziściła się sytuacja przyjętych założeń, sprawa jest bardziej skomplikowana. Wówczas, gdy $\bar{r} = r_g$, przypadek jest trywialny i nie ma problemu wygładzania szeregu czasowego, zachodzi bowiem *casus* wzrostu równomiernego, a więc $(r_g = \bar{r}) = f(t_i) = \text{const}$. Znalezienie teoretycznego $\bar{Y}_0$ spełniającego wymóg rzeczywistego Y_n według stopy wzrostu $\bar{r}$ oznacza spełnienie równania:

$$\bar{Y}_0 = \frac{Y_n}{(1+\bar{r})^n} = \frac{\bar{Y}_0(1+\bar{r})^n}{(1+\bar{r})^n} = \bar{Y}_0$$

1. Przypadek $\bar{r} < r_g$, czyli $\bar{Y}_n < Y_n$



Rys. 3.

Wyrównanie dla Y_n i $\bar{r}$ przy $\bar{r} < r_g$

Źródło: koncepcja własna.

W tej sytuacji różnica między sumami dla $t = n$ i $t = n - 1$ w *Tablicach* będzie mniejsza od Y_n i dlatego musimy obliczać różnice między $t = n + 1$ a $t = n$, a jeżeli i ta różnica będzie mniejsza od Y_n, to postępujemy kolejno dla $t = n + 2$ i $t = n + 1$ itd., aż do znalezienia różnicy równej wyrazowi rzeczywistemu Y_n.

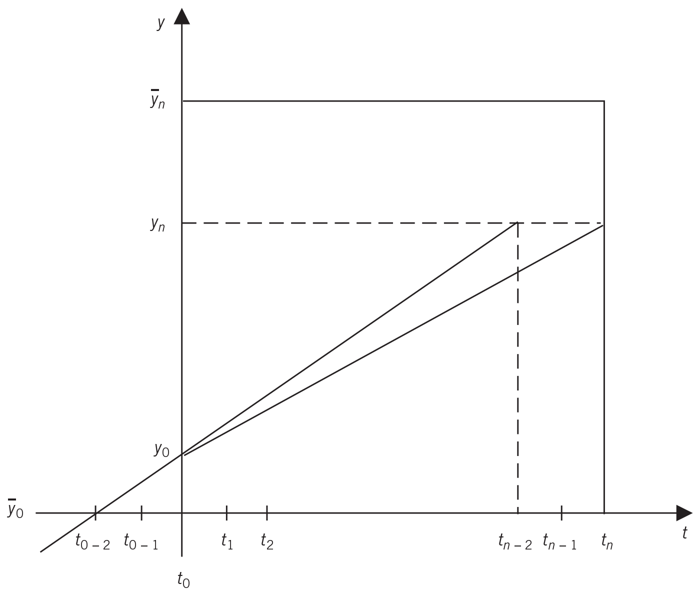
Dokonując dla n okresów (momentów) deflacji według wartości liczbowej deflatora $(1 + \bar{r})$, uzyskamy ciąg wyrazów

$$Y_n:(1+\bar{r}), Y_n:(1+\bar{r})^2, \ldots, Y_n:(1+\bar{r})^{n-1}, Y_n:(1+\bar{r})^n$$

czyli wartości liczbowe szeregu teoretycznego łącznie z wyrazem podstawowym $\overline{Y}_0 = Y_n:(1+\bar{r})^n$, który z definicji będzie większy od rzeczywistego, czyli $Y_0 < \overline{Y}_0$. Tę procedurę postępowania obrazuje rysunek 3.

2. Przypadek $\bar{r} > r_g$, czyli $\overline{Y}_n > Y_n$

W tej sytuacji różnica między sumami dla $t = n$ i $t = n - 1$ w *Tablicach* będzie większa od Y_n i dlatego musimy cofać się, licząc kolejno różnice dla $t = n - 1$ i $t = n - 2$ itd., aż do uzyskania różnicy równej wyrazowi rzeczywistemu Y_n. Jednakże w tym przypadku będziemy musieli cofnąć się do ujemnych wartości czasu, a więc $t = 0 - 1$, $t = 0 - 2$... Jednakże *Tablice* rozpoczynają od $t = 0$ i dlatego „w ujemnym czasie" będziemy deflowali deflatorem $(1 + \bar{r})$ liczbę 1,000, podnosząc do potęgi 1, 2, ..., tzn. do momentu, gdy liczba okresów zrówna się z liczbą n. W tym przypadku będzie relacja $\overline{Y}_0 < Y_0$. Przedstawmy to na rysunku.



Rys. 4.

Wyrównanie dla Y_n i $\bar{r}$ przy $\bar{r} > r_g$
Źródło: koncepcja własna.

VII

W uzupełnieniu rozważań teoretycznych i przede wszystkim dla ilustracji wyżej rozważanych kwestii co do „dobroci" dopasowania danych teoretycznych do danych rzeczywistych przedstawimy konkretny przykład. Tyczy się on dynamiki średnich miesięcznych emerytur i rent rolników indywidualnych brutto w 2003 r., przy przyjęciu ich poziomu w grudniu 2002 r. za 100. Wówczas obliczone przeciętne tempo zmian tych miesięcznych świadczeń społecznych wynosiło: r_g = 0,46% i $\bar{r}$ = 0,80%. Niżej przedstawiamy szereg faktyczny oraz szeregi teoretyczne (wyrównane dane) według metody r_g i $\bar{r}$.

Tabela 1.

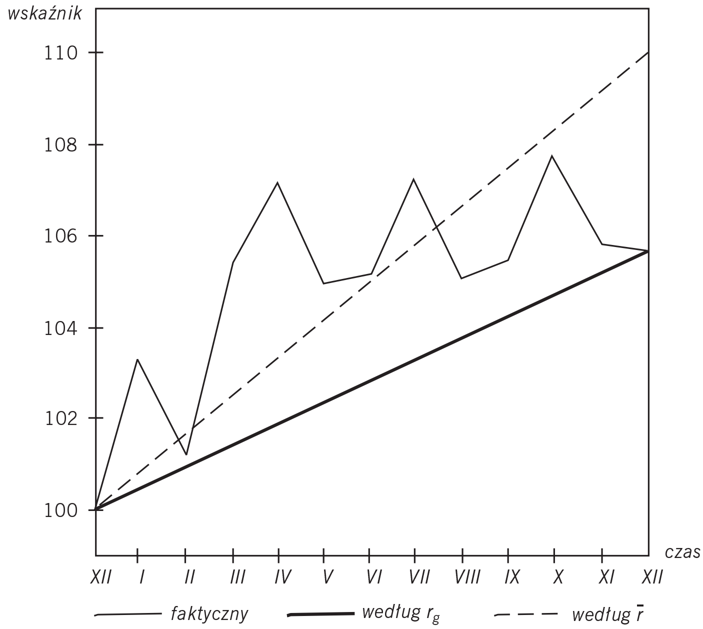
Szeregi faktyczny i teoretyczne wyrównane metodami r_g i $\bar{r}$ przeciętnych miesięcznych emerytur i rent brutto rolników indywidualnych w 2003 r. (grudzień 2002 r. = 100)

Miesiące	Szereg faktyczny	Szereg teoretyczny według	
		r_g	$\bar{r}$
I	103,3	100,5	100,8
II	101,3	100,9	101,6
III	105,4	101,4	102,4
IV	107,2	101,9	103,2
V	105,0	102,3	104,1
VI	105,1	102,8	104,9
VII	107,2	103,3	105,7
VIII	105,1	103,7	106,6
IX	105,5	104,2	107,4
X	107,7	104,7	108,3
XI	105,9	105,2	109,2
XII	105,7	105,7	110,0

Źródło: na podstawie: „Biuletyn Statystyczny", nr 1, luty 2004 r., GUS, tabl. 16, s. 60; obliczenia własne.

W przedstawionym przykładzie kwadrat odchyleń danych teoretycznych od danych rzeczywistych wynosi: dla wyrównania wedle metody r_g = 93,02, a dla $\bar{r}$ = 83,04. A więc metoda $\bar{r}$ czyni minimalizację kwadratów odchyleń danych faktycznych od danych teoretycznych w porównaniu z metodą r_g. Czyni więc zadość wymaganiom metody najmniejszych kwadratów.

Wyraziście przedstawione wartości liczbowe szeregów faktycznego i teoretycznych obrazuje rysunek 5.



Rys. 5.

Szeregi faktyczny i teoretyczne według r_g i $\bar{r}$ dynamiki przeciętnych miesięcznych emerytur i rent rolników indywidualnych brutto w 2003 r. (grudzień 2002 r. = 100)

Źródło: tabela 1.

Rysunek 5. nie wymaga komentarza. Zaniżanie średniej dynamiki przez metodę r_g (nieważonej średniej geometrycznej) jest widoczne, jak na przysłowiowej dłoni.

*

Przedstawiony wyżej wywód czyni w pełni uzasadnienie użycia w stosunku do metody $\bar{r}$ terminu: substytut metody najmniejszych kwadratów.

Autor z zadowoleniem odnotowuje, że metoda $\bar{r}$ uzyskuje coraz bardziej ugruntowane „prawa obywatelskie". Świadczą o tym powołania się na nią w publikacjach naukowych i podręcznikach.

Omawiana tu właściwość metody $\bar{r}$ (wykorzystującej sumę wyrazów faktycznego szeregu chronologicznego) poza jej warstwą ściśle naukową spełnia także funkcję w pracy dydaktycznej związanej ze statystyką i ekonometrią.

Abstract

The Method $\bar{r}$ as substitute for the Least Square Method

The subject of this paper was presentation of the properties of the Method $\bar{r}$, i.e. a method utilizing the sum of the terms of a real time series in calculating the average (mean) of the growth pace (rate). This property permits to use the Method $\bar{r}$ and its tables as substitute for the Least Square Method in fitting the time series wherever an exponentialtype function is to be adjusted to real (experimental) data. The solution of the problem was presented in two variants: a) real initial (basic) term and real sum of terms of the time series in question; 2) real final term and real sum of terms of the time series in question.