Trend
Metody analizy szeregów czasowych

Modele ze zmienną czasową

Zjawiska ekonomiczne czy społeczne kształtowane są przez szereg czynników, których natura często jest na tyle skomplikowana, że trudno jest uchwycić je w modelach typu regresyjnego.

Czasem możliwe staje się zastąpienie takich modeli modelami dynamicznymi, w których rolę zmiennej objaśniającej pełni czas.

Oczywiście czas z reguły nie będzie przyczyną analizowanego zjawiska, ale będzie reprezentantem zbioru zmiennych niezależnych kształtujących to zjawisko.

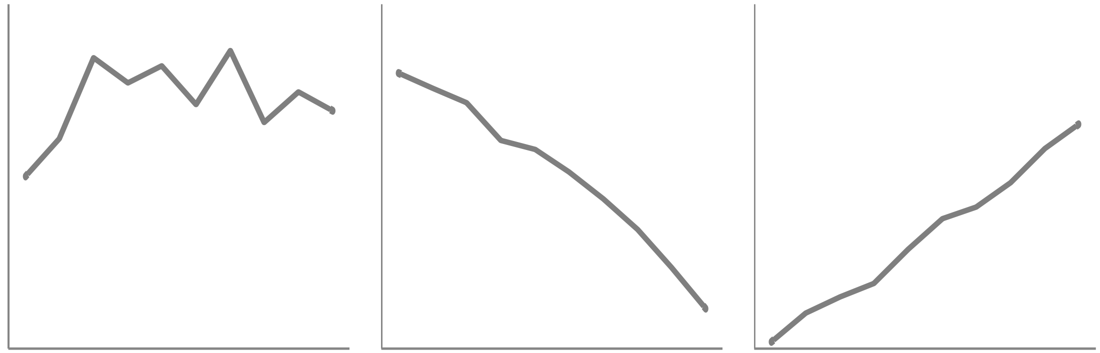
Formalnie oznacza to, że zamiast analizy modelu typu $\boldsymbol{y=f(x_1,x_2,\dots,x_k)}$

analizuje się modele postaci $\boldsymbol{y=f(t)}$.

Trend w szeregach czasowych

Główne przyczyny odpowiedzialne za tendencję rozwojową zmiennych ekonomicznych działają w sposób trwały. Tworzą one długoterminowy kierunek zmian. W zależności od tego kierunku mówi się o **trendzie wzrostowym, spadkowym bądź bocznym** (stałym).

Rysunek:



Trend boczny Trend spadkowy Trend wzrostowy

Odchylenia okresowe

Przyczyny okresowe powodują powstawanie regularnych zmian wartości badanego zjawiska ekonomicznego wokół tendencji rozwojowej. Ich cechą charakterystyczną jest powrót badanego zjawisko do stanu wyjściowego po zakończeniu pełnego cyklu. Wyróżniamy odmiany wahań okresowych:

- **wahania koniunkturalne** (fluktuacyjne) będące wahaniami o okresach od kilkuletnich poprzez kilkunastoletnie do cykli długich. W literaturze przedmiotu opisywane są cykle Kitchina, Juglara, Kuznetsa, Kondratiewa czy Wagernana. Przyczyny wahań cyklicznych zmiennych ekonomicznych mogą być zarówno natury ekonomicznej (np. zmiany stanu zapasów, polityka fiskalna i pieniężna i inne) jak i pozaekonomicznej (siły natury, cykle polityczne i inne). Cechą charakterystyczną takich cykli jest to, że dotyczą one zmian wieloletnich i z reguły charakteryzują się przeciętną powtarzalnością,

Odchylenia okresowe

- **wahania specjalne,** to wahania dotykające wybrane dziedziny życia gospodarczego. Najczęściej przedstawianym przykładem takich wahań są tzw. „świńskie górki” w produkcji żywca wieprzowego;
- **wahania sezonowe**, to wahania, które ujawniają się w okresach krótszych niż rok, tj. w okresach miesięcznych lub kwartalnych.

W odróżnieniu od wahań koniunkturalnych i specjalnych charakteryzują się one dość dużą powtarzalnością, tj. czas trwania takich cykli ograniczony jest dokładnie do roku.

- **wahania krótkookresowe**, to wahania o naturze zbliżonej do

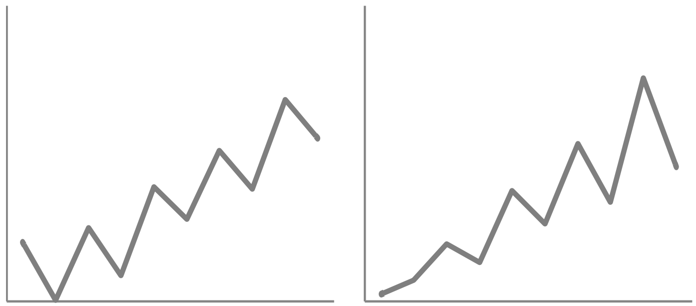
Odchylenia przypadkowe

Ostatnią grupą przyczyn wpływających na kształtowanie szeregów czasowych są **przyczyny przypadkowe**, które powodują nieprzewidywalne zmiany kierunku i siły analizowanego zjawiska. Wahania przypadkowe to zarówno wahania o charakterze katastrofalnym, takie jak klęski żywiołowe czy wojny, jak i wahania powodowane słabymi przyczynami ubocznymi, które jednak w swojej masie powodują właśnie tę nieprzewidywalność zmian wartości szeregu czasowego.

Analiza kształtowania szeregów czasowych

Badane zjawisko ze względu na dynamikę wahań okresowych można rozłożyć na składowe według **formuły addytywnej** lub **multiplikatywnej**.

Procedura rozkładu szeregu czasowego na składowe to **dekompozycja.**



Rysunek. Zjawisko rozwijające się addytywnie (a) i multiplikatywnie (b).

Analiza kształtowania szeregów czasowych

W przypadku **formuły addytywnej** badane zjawisko Y_t jest sumą składowych:

$$\boldsymbol{Y_t = T_t + C_t + S_t + I_t}$$

gdzie:

T_t – trend;

C_t – składnik cykliczny;

S_t – składnik sezonowy;

I_t – składnik nieregularny

Analiza kształtowania szeregów czasowych

Wahania addytywne mają stałą amplitudę wahań okresowych (cyklicznych i sezonowych), która nie zależy od poziomu zjawiska w czasie. Jest ona wyrażana w ujęciu absolutnym, w tych samych jednostkach, w których mierzona jest cecha *Y*. Nakłada się ona na trend addytywnie, czyli **wartość trendu jest regularnie pomniejszana lub powiększana o pewne stałe wartości będące obrazem wahań okresowych.** Wszystkie składniki wyrażone są w tych samych jednostkach co zmienna *Y*.

Analiza kształtowania szeregów czasowych

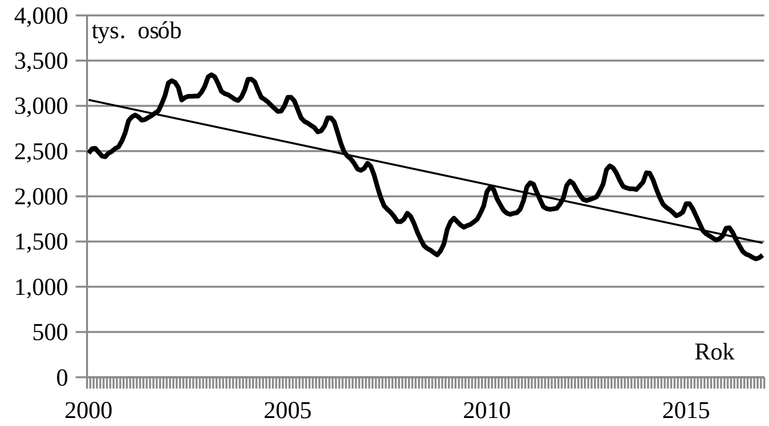
W przypadku **formuły multiplikatywnej** badane zjawisko składa się za pomocą iloczynu czynników:

$$Y_t = T_t \cdot C_t \cdot S_t \cdot I_t.$$

W tej formule składnik obrazujący linię trendu T_t jest wyrażony w takich samych jednostkach, jak wyrazy szeregu Y, a pozostałe składniki mają postać **wskaźników.** Formułę multiplikatywną stosuje się do zjawisk, w których amplituda zależy od rozwoju zjawiska w czasie, co przejawia się tym, **że wartości trendu są proporcjonalnie powiększane bądź pomniejszane.** Inaczej można powiedzieć, że

Analiza kształtowania szeregów czasowych

Na rysunku przedstawiono kształtowanie liczby bezrobotnych w Polsce w latach 2000-2016. Widoczne są składowe szeregów czasowych, jest obecny malejący trend, dwa cykle koniunkturalne, roczne wahania sezonowe oraz wahania przypadkowe.

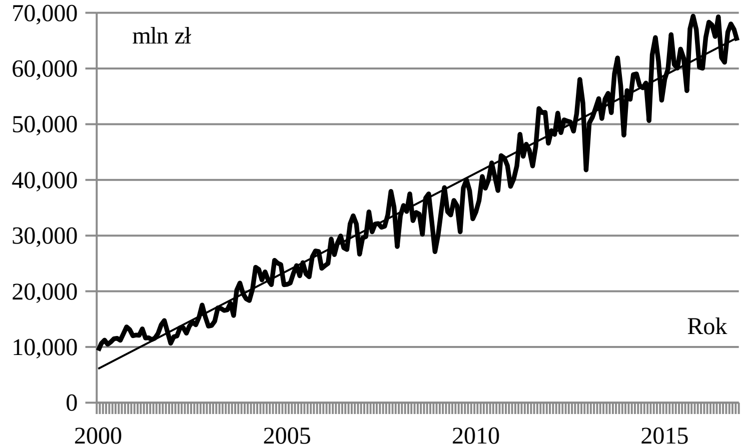


Rysunek. Bezrobotni zarejestrowani w Polsce w latach 2000-2016 (stan na koniec miesiąca).

Źródło: Główny Urząd Statystyczny.

Analiza kształtowania szeregów czasowych

Rysunek. Eksport towarów w cenach bieżących w Polsce w latach 2000-2016. Źródło: Główny Urząd Statystyczny.



Powyższy wykres przedstawiający eksport towarów w Polsce jest przykładem trendu wzrostowego, ale już obecność cykli koniunkturalnych nie jest pewna. Wprawdzie w latach 2008-2009 miało miejsce pewne spowolnienie eksportu, i okres ten można byłoby uznać za rozdzielający szereg na dwa podokresy. Najbardziej prawdopodobne jest występowanie dość silnych wahań

Metody wyodrębniania trendu

Do najprostszych metod wyodrębnienia trendu z szeregu czasowego można zaliczyć **metodę analityczną**, w której trend wyznaczany jest za pomocą funkcji trendu (np. liniowej, wykładniczej, wielomianowej lub innej) oraz **metodę mechaniczną**, w której trend wyznaczany jest za pomocą średniej ruchomej.

Metody wyodrębniania trendu

Prostoliniową funkcję trendu przedstawia się w postaci równania:

$$\boldsymbol{y_t = \square_0 + \square_1 t,}$$

gdzie:

$\square_0$ – wyraz wolny równania;

$\square_1$ – współczynnik trendu

Wyraz wolny $\square_0$ wyznacza teoretyczną wartość zmiennej w okresie oznaczonym przez t=0.

Natomiast współczynnik regresji $\square_1$ wskazuje na średni przyrost wartości ocenianej zmiennej w jednostce czasu.

Metody wyodrębniania trendu

Sytuacja jest tutaj analogiczna do prostoliniowego równania regresji typu:

$y_t = \alpha_0 + \alpha_1 x$, przy czym rolę zmiennej objaśniającej pełni tutaj czas.

Parametry liniowego równania trendu można, podobnie jak równania regresji, wyznaczyć przy użyciu metody najmniejszych kwadratów (MNK), rozwiązując układ równań:

$$\begin{cases} \sum y_t = n\alpha_0 + \alpha_1 \sum t \\ \sum y_t t = \alpha_0 \sum t + \alpha_1 \sum t^2 \end{cases}$$

Równanie trendu liniowego pokazuje długookresowy kierunek, w którym podąża badane zjawisko. Jest to model odpowiedni, gdy w analizowanym zjawisku obserwuje się średnio stały przyrost bezwzględny wartości ocenianej zmiennej.

Modele trendu

- W niektórych sytuacjach zastosowanie modelu liniowego nie będzie skuteczne. Szczególnie w sytuacji, gdy zjawisko zmienia się średnio nie o stałą wartość, a średnio o stały iloraz. W takich sytuacjach stosujemy model wykładniczy:

$$y_t = \square_0 \square_1^t,$$

który poprzez obustronne zlogarytmowanie, można sprawdzić do postaci liniowej:

$$\ln y_t = \ln \square_0 + t \square \ln \square_1$$

Parametr $\square_0$ wyznacza teoretyczną wartość zmiennej w okresie oznaczonym przez ***t*=0:**

$$y_0 = \square_0 \square_1^0 = \square_0 \square 1 = \square_0.$$

Parametr $\square_1$ wyznacza **średni względny przyrost** wartość zmiennej w jednostce czasu, może być on nazwany **średnim tempem wzrostu**:

Modele trendu

Parametry tak zlinearyzowanego równania wyznacza się już przy użyciu metody MNK rozwiązując układ równań, z tym, że zamiast zmiennej ***y*** używa się zmiennej **ln*y*,** a zamiast parametrów α_0 i α_1, otrzymuje się parametry **ln**α_0 i **ln**α_1**,** które poprzez zastosowanie funkcji ekspotencjalnej można doprowadzić do postaci α_0 i α_1**.**

$$\begin{cases} \sum_{t=1}^{n} \ln y_t = n \ln a_0 + \ln a_1 \sum_{t=1}^{n} t \\ \sum_{t=1}^{n} \ln y_t t = \ln a_0 \sum_{t=1}^{n} t + \ln a_1 \sum_{t=1}^{n} t^2 \end{cases}$$

Modele trendu

Parametry $\boldsymbol{ln\square_0}$ i $\boldsymbol{ln\square_1}$ poprzez zastosowanie funkcji ekspotencjalnej można doprowadzić do postaci $\square_0$ i $\square_1$.

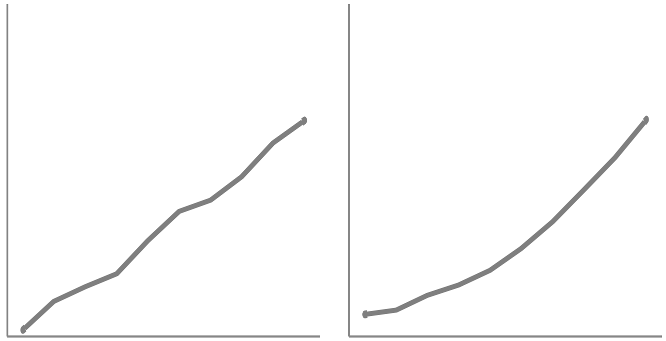
Jeżeli problem dotyczy jedynie określenia średniego tempa wzrostu, to wykorzystując właściwość, że:

$$\mathbf{1 + \boldsymbol{a} \;\square\; e^{\boldsymbol{a}}, \text{ dla } \boldsymbol{a}\square 0,}$$

współczynnik otrzymany w modelu zlinearyzowanym można bezpośrednio interpretować jako poszukiwane średnie tempo wzrostu.

Modele trendu

W ocenie jakości uzyskanego równania stosuje się te same narzędzia jak w przypadku równania regresji, czyli współczynnik determinacji, współczynnik zbieżności, odchylenie standardowe składnika resztowego czy średni błąd względny.



Rysunek. Linowy i wykładniczy model trendu.

Ocena modelu

Ocenę dopasowania modelu do danych empirycznych można przeprowadzić na podstawie **współczynnika determinacji:**

$$R^2 = \frac{\sum_{i=1}^{n} (y_i' - \bar{y})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}$$

gdzie:

y_i – wartości empiryczne zmiennej Y;

y_i' – wartości teoretyczne zmiennej Y wyznaczone z równania dla zaobserwowanych wartości empirycznych x_i zmiennej X;

$\bar{y}$ – przeciętny poziom wartości zmiennej Y.

Ocena modelu

- **Współczynnik zbieżności**

$$\varphi^2 = \frac{\sum_{i=1}^{n} (y_i - y_i')^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}$$

Określa część zmienności zmiennej Y niewyjaśnioną zmiennością zmiennej X. Zatem:

Dokładność szacunków

Dokładność takich szacunków można ocenić za pomocą:

1. **Wariancji składnika resztowego** (nie ma interpretacji)

$$S_e^2 = \frac{1}{n-2}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \frac{1}{n-2}\sum_{i=1}^{n} e_i^2$$

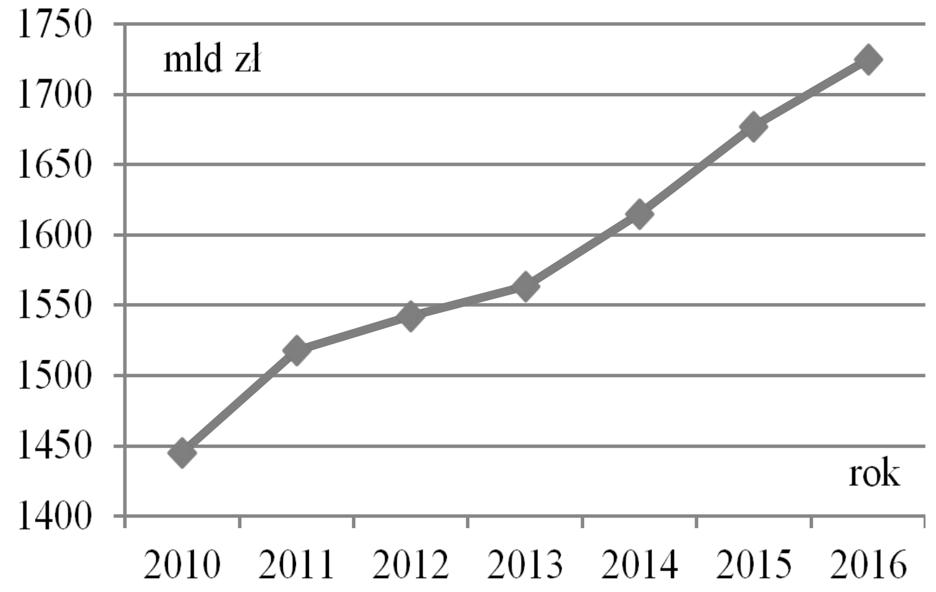
2. **Odchylenia standardowego składnika resztowego** (inaczej średni błąd oszacowania). Informacje o ile wartości teoretyczne różnią się do empirycznych. $S_e = \sqrt{S_e^2}$

3. **Średniego błędu względnego**. Informuje o skali błędu oszacowania w stosunku do wartości średniej zmiennej

$$V_e = \frac{S_e}{\bar{y}} \cdot 100\%$$

Zadanie 1. Dane dotyczą Produktu Krajowego Brutto Polski w latach 2010-2016. PKB w kolejnych latach przedstawiono w cenach z roku 2010.

rok	2010	2011	2012	2013	2014	2015	2016
PKB (w mld zł)	1445,30	1517,81	1542,22	1563,68	1615,02	1677,11	1725,15



Rys. 1. Kształtowanie PKB Polski w latach 2010-2016

Dla oszacowania liniowej funkcji trendu $y_t = \alpha_0 + \alpha_1 t$ należy rozwiązać układ równań normalnych. Obliczenia pomocnicze pokazano w tabeli:

rok	t	y_t	$y_t \cdot t$	t^2
2010	1	1445,30	1445,30	1
2011	2	1517,81	3035,62	4
2012	3	1542,22	4626,66	9
2013	4	1563,68	6254,72	16
2014	5	1615,02	8075,10	25
2015	6	1677,11	10062,66	36
2016	7	1725,15	12076,05	49
	28	11086,29	45576,11	140

Rozwiązując układ:

$$\begin{cases} \sum_{t=1}^{n} y_t = n\alpha_0 + \alpha_1 \sum_{t=1}^{n} t \\ \sum_{t=1}^{n} y_t t = \alpha_0 \sum_{t=1}^{n} t + \alpha_1 \sum_{t=1}^{n} t^2 \end{cases}$$

$$11\,086{,}29 = 7\alpha_0 + 28\alpha_1$$

$$45\,576{,}11 = 28\alpha_0 + 140\alpha_1$$

uzyskuje się liniową funkcję trendu postaci:

$$y' = 1407{,}91 + 43{,}96t$$

Uzyskana wartość współczynnika trendu oznacza, że z roku na rok wartość PKB rosła średnio o 43,96 mld zł (w cenach z roku 2010).

Ocenę jakości wyznaczonego równania przeprowadzono przy użyciu współczynnika determinacji, odchylenia standardowego składnika resztowego oraz średniego błędu względnego. Obliczenia pomocnicze zawarto w tabeli:

rok	t	y_t	y'=1407,91 + 43,96t	$(y_t' - \bar{y})^2$	$(y_t - \bar{y})^2$	$(y_t - y_t')^2$
2010	1	1445,30	1451,87	17394,97	19171,17	43,1649
2011	2	1517,81	1495,83	7731,68	4349,40	483,1204
2012	3	1542,22	1539,79	1933,36	1725,57	5,9049
2013	4	1563,68	1583,75	0,00	403,21	402,8049
2014	5	1615,02	1627,71	1931,60	977,19	161,0361
2015	6	1677,11	1671,67	7728,17	8714,22	29,5936
2016	7	1725,15	1715,63	17389,70	19991,13	90,6304
	28	11086,29		54109,49	55331,89	1216,26

$$\bar{y} = \frac{11086,29}{7} = 1583,76$$

Model liniowy dobrze oddaje kształtowanie się ocenianego zjawiska, gdyż charakteryzuje się wysokim współczynnikiem determinacji:

$$R^2 = \frac{\sum_{t=1}^{n} (\hat{y}_t - \bar{y})^2}{\sum_{t=1}^{n} (y_t - \bar{y})^2} = \frac{[illegible]}{[illegible]} = [illegible]$$

oraz niską wartością odchylenia standardowego składnika resztowego:

$$S_e = \sqrt{\frac{1}{n-2} \sum_{t=1}^{n} (y_t - \hat{y}_t)^2} = \sqrt{\frac{1}{T-2} [illegible]} = [illegible]$$

i średniego błędu względnego:

$$V_e = \frac{S_e}{\bar{y}} \cdot 100\% = \frac{[illegible]}{[illegible]} \cdot 100\% = [illegible].$$

Do danych dotyczących Produktu Krajowego Brutto można także zastosować **model wykładniczy**

$$y_t = \square_0 \square_1^t$$

Aby zastosować układ równań normalnych do oszacowania parametrów modelu wykładniczego należy uprzednio wartości szeregu zlogarytmować, a następnie oszacować parametry modelu zlinearyzowanego

$$\ln y_t = \ln\square_0 + t\square\ln\square_1$$

Obliczenia pomocnicze konieczne do zastosowania układu pokazano w tabeli:

rok	t	y	ln$y$	ln$y \cdot t$	t^2
2010	1	1445,30	7,2761	7,28	1
2011	2	1517,81	7,3250	14,65	4
2012	3	1542,22	7,3410	22,02	9
2013	4	1563,68	7,3548	29,42	16
2014	5	1615,02	7,3871	36,94	25
2015	6	1677,11	7,4248	44,55	36
2016	7	1725,15	7,4531	52,17	49
	28		51,5619	207,0243	140

Zatem rozwiązując układ równań:

$$\begin{cases} \sum_{t=1}^{n} \ln y_t = n \ln \alpha_0 + \ln \alpha_1 \sum_{t=1}^{n} t \\ \sum_{t=1}^{n} \ln y_t t = \ln \alpha_0 \sum_{t=1}^{n} t + \ln \alpha_1 \sum_{t=1}^{n} t^2 \end{cases}$$

$$51{,}5619 = 7 \cdot \ln \alpha_0 + 28 \cdot \ln \alpha_1$$

$$207{,}0243 = 28 \cdot \ln \alpha_0 + 140 \cdot \ln \alpha_1$$

uzyskuje się liniową funkcję trendu postaci:

$$\ln y' = 7{,}255 + 0{,}0277t$$

$$\ln y' = 7{,}255 + 0{,}0277t$$

Uzyskana wartość współczynnika trendu oznacza, że z roku na rok wartość PKB rosła średnio o 2,77%. Wartość ta jest wartością przybliżoną, gdyż po odlinearyzowaniu powyższych parametrów uzyskuje się właściwe parametry modelu wykładniczego:

$$\square_0 = e^{7,255} = 1\,415{,}16$$

$$\square_1 = e^{0,0277} = 1{,}0281$$

a więc model postaci:

$$y' = 1415{,}16 \square 1{,}0281^t$$

Z czego wynika, że w kolejnych latach wartość PKB była średnio 1,0281 razy większa niż rok wcześniej, lub inaczej, że średnioroczne tempo wzrostu PKB wyniosło 2,81%. Zatem przybliżenie współczynnika trendu przed odlinearyzowaniem jest dość dobre.

Warto sprawdzić, że parametry jakościowe modelu wykładniczego, podobnie jak modelu liniowego, są bardzo dobre. Pamiętać należy, że wzory dotyczące współczynnika determinacji, odchylenia standardowego składnika resztowego oraz średniego błędu względnego należy zastosować do modelu zlinearyzowanego, a więc do wartości empirycznych ln*y* oraz modelu ln*y*' = 7,255 + 0,0277*t*.