

12. Modele Dynamiczne

Modele dynamiczne służą do modelowania procesów, dla których reakcja zmiennej zależnej na błędy losowe lub zmiany zmiennych niezależnych jest rozciągnięta w czasie. Modele takie estymujemy przy użyciu szeregów czasowych.

Będziemy zajmowali się klasą modeli, które można przedstawić w kategoriach wielomianów operatora opóźnień

$$A(L)y_t = B(L)x_t + C(L)\varepsilon_t \quad (12.1)$$

Modele o takiej postaci nazywamy modelami *ARMAX*. W modelu *ARMAX* można wyróżnić część autoregresyjną *AR* (**A**uto**R**egressive), związaną z opóźnionymi y_t , część związaną ze średnią ruchomą (**M**oving **A**verage) błędów losowych, i część związaną z opóźnionymi x_t nazywaną częścią *DL* (**D**istributed **L**ag).

12.1. Modele *DL*

Model, w którym występuje jedynie część *DL* jest szczególnie prosty w analizie ponieważ jeśli x_t jest w nim nielosowe to tym samym opóźnione x_t są także nielosowe i model spełnia założenia *KMRL*. Szczególnie prosty przypadek zachodzi, gdy x_t zawiera tylko stałą i jedną zmienną:

$$y_t = \mu + \beta_0 x_t + \dots + \beta_p x_{t-p} + \varepsilon_t \quad (12.2)$$

W takim modelu natychmiastowa reakcja y_t na zmianę x_t o jednostkę na wynosi β_0 . Jeśli x_t pozostanie na tym wyższym poziomie także w następnych okresach, to w długim okresie wpływ zmiany x_t na y_t będzie równy $\sum_{i=1}^s \beta_i$. Współczynnik β_0 nazywany jest mnożnikiem bezpośrednim a $\beta = \sum_{i=1}^s \beta_i$ mnożnikiem skumulowanym. Można przeformułować model tak, by parametrami były udziały poszczególnych opóźnionych efektów w efekcie całkowitym

$$y_t = \mu + \beta \sum_{i=0}^s w_i x_{t-i} + \varepsilon_t$$

gdzie $w_i = \frac{\beta_i}{\beta}$ nazywamy wagami opóźnień. Średnie opóźnienie zmiany y_t powstałej na skutek zmiany x_t wynosi $\bar{w} = \sum_{i=0}^s i w_i$. Podane definicje można łatwo uogólnić na przypadek, kiedy x_t zawiera więcej niż jedną zmienną definiując mnożniki, wagi i średnie opóźnienie dla każdego z elementów wektora x_t z osobna.

Problemem przy estymacji modeli *DL* jest ustalenie ilości opóźnień, które powinny być użyte w modelu. Nie powinno się stosować zbyt dużej ilości opóźnień, ponieważ wiąże się to z utratą dużej ilości stopni swobody. Jeśli jednak w modelu użyjemy jednak zbyt małej ilości opóźnień, to pojawi się problem zmiennych pominiętych. Do ustalenia ilości opóźnień można posłużyć się testem *F* lub jednym z kryteriów informacyjnych. Prawidłowa procedura ustalania ilości opóźnień opiera się na testowaniu łącznej niestotności najbardziej odległych opóźnień. Na początku ustalamy s^* - maksymalną liczbę opóźnień w modelu. Następnie testujemy hipotezę $H_0 : \beta_{s^*} = 0$. Jeśli nie ma podstaw do jej odrzucenia, testujemy hipotezę łączną $H_0 : \beta_{s^*} = \beta_{s^*-1} = 0$. Testowanie to kontynuujemy tak długo aż znajdziemy takie s , dla którego hipoteza zerowa $H_0 : \beta_{s^*} = \dots = \beta_s = 0$ zostanie odrzucona. Uzyskane s przyjmujemy jako ilość opóźnień w modelu. W podobny sposób można posłużyć się kryteriami informacyjnymi, szukając takiego s , dla którego kryteria mają minimalną wartość. Metoda ta jest szczególnym przypadkiem metody od ogólnego do szczegółowego.

Innym sposobem na ograniczenie liczby traconych stopni swobody jest złożenie, że przyjęcie założenia, że rozkład opóźnień dany jest pewną funkcją o niewielkiej ilości parametrów. Dwie najpopularniejsze formy przyjmowane dla rozkładu opóźnień to opóźnienia wielomianowe Almon i opóźnienia geometryczne Koycka.

• Opóźnienia Almon

Modele wielomianowych opóźnień *PDL* (**P**olynomial **D**istributed **L**ags) są to modele *DL* z rozkładem opóźnień mającym formę wielomianu. W modelu tym zakładamy, że prawdziwe są następujące ograniczenia dotyczące parametrów β_i w modelu

$$\beta_i = \alpha_0 + \alpha_1 i + \alpha_2 i^2 + \dots + \alpha_q i^q \quad i = 1, \dots, p \quad q \leq p \quad (12.3)$$

Ograniczenia te łatwo wprowadzić jest bezpośrednio do modelu (12.2)

$$\begin{aligned}
 y_t &= \mu + \beta_0 x_t + \beta_1 x_{t-1} + \dots + \beta_p x_{t-p} + \varepsilon_t \\
 &= \mu + \beta_0 x_t \\
 &\quad + (\alpha_0 + \alpha_1 + \alpha_2 + \dots + \alpha_q) x_{t-1} \\
 &\quad + (\alpha_0 + 2\alpha_1 + 4\alpha_2 + \dots + 2^q \alpha_q) x_{t-2} \\
 &\quad \vdots \\
 &\quad + (\alpha_0 + p\alpha_1 + p^2\alpha_2 + \dots + p^q \alpha_q) x_{t-p} \\
 &= \mu + \beta_0 x_t \\
 &\quad + \alpha_0 (x_{t-1} + x_{t-2} + \dots + x_{t-p}) \\
 &\quad + \alpha_1 (x_{t-1} + 2x_{t-2} + \dots + px_{t-p}) \\
 &\quad + \alpha_1 (x_{t-1} + 4x_{t-2} + \dots + p^2 x_{t-p}) \\
 &\quad \vdots \\
 &\quad + \alpha_q (x_{t-1} + 2^q x_{t-2} + \dots + p^q x_{t-p}) \\
 &= \mu + \beta_0 x_t + \alpha_0 z_{0t} + \dots + \alpha_p z_{qt} + \varepsilon_t
 \end{aligned}$$

przy czym w ostatnim równaniu

$$z_{kt} = \sum_{j=1}^p j^k x_{t-j}, \text{ dla } k = 0, \dots, q$$

Taki przeformułowany model można bez trudu wyestymować za pomocą *MNK*. Wielkości współczynników β_0, \dots, β_p można uzyskać jako kombinację liniową wyestymowanych parametrów α_j korzystając ze wzoru (12.3).

Stosując metodę natrafiamy na podobny problem jak w przypadku ustalania ilości opóźnień, ponieważ musimy ustalić rząd wielomianu użytego do aproksymacji opóźnień wielomianu. Możemy to zrobić stosując podobną do opisaną wcześniej procedurę sekwencyjnego testowania istotności największych potęg wielomianu.

- **Opóźnienia geometryczne (Koycka)**

Innym popularnym modelem opóźnień są opóźnienia geometryczne. Załóżmy, że wiadomo, że wpływ zmiany w zmiennej niezależnej ma stopniowo zanikający wpływ na zmienną zależną, ale nie można powiedzieć, po ilu okresach wpływ ten ustaje całkowicie. W takim przypadku można założyć, że wpływ kolejnych zmiennych opóźnionych geometrycznie maleje i $\beta_{i+1} = \lambda\beta_i$, gdzie $|\lambda| < 1$ i $\beta_0 = \beta(1 - \lambda)$. Model (12.2) przyjmuje w tym przypadku postać

$$\begin{aligned} y_t &= \mu + \beta \sum_{i=0}^{\infty} (1 - \lambda) \lambda^i x_{t-i} + \varepsilon_t \\ &= \mu + \beta(1 - \lambda) \sum_{i=0}^{\infty} (\lambda L)^i x_t + \varepsilon_t \\ &= \mu + \beta \frac{1 - \lambda}{1 - \lambda L} x_t + \varepsilon_t \end{aligned} \tag{12.4}$$

Modelu w tej postaci nie można wyestymować, ponieważ zawiera on nieskończoną ilość zmiennych niezależnych. Jeśli jednak pomnożymy obie strony (12.4) przez $1 - \lambda L$, to po przeniesieniu y_{t-1} na prawą stronę, otrzymamy

$$y_t = \mu(1 - \lambda) + \lambda y_{t-1} + \beta(1 - \lambda)x_t + (1 - \lambda L)\varepsilon_t \tag{12.5}$$

Parametrów takiego modelu nie można wyestymować za pomocą *MNKG*, ponieważ występuje w nim problem równocześnieści, skoro błąd losowy $u_t = (1 - \lambda L)\varepsilon_t = \varepsilon_t - \lambda\varepsilon_{t-1}$ jest skorelowany

ze zmienną niezależną y_{t-1} . Prawidłowym sposobem estymacji równania 12.5 jest zastosowanie estymatora MZI przy czym za instrumenty można przyjąć zmienne x_{t-1}, \dots, x_{t-s} .

Przykład 12.1 *Oczekiwania adaptacyjne*

Estymujemy następującą funkcję konsumpcji

$$C_t = a + bY_t^* + \varepsilon_t \quad (12.6)$$

gdzie Y_t^ jest oczekiwanym długookresowym dochodem. Załóżmy, że oczekiwania są kształtowane w następujący sposób*

$$\Delta Y_t^* = (1 - \lambda) (Y_t - Y_{t-1}^*)$$

Zmiana oczekiwań co do długookresowego dochodu jest w tym modelu proporcjonalna do różnicy między rzeczywistym i oczekiwanym dochodem. Wzór ten można zapisać jako

$$Y_t^* = \lambda Y_{t-1}^* + (1 - \lambda) Y_t$$

co oznacza, że oczekiwany długookresowy dochód jest średnią ważoną dochodu rzeczywistego i poziomu dochodu jaki oczekiwaliśmy w przeszłości. Zapisując ten wzór przy użyciu operatora opóźnień stwierdzamy, że

$$(1 - \lambda L) Y_t^* = (1 - \lambda) Y_t$$

Oczekiwania konsumentów nie są obserwowalne, jednak

$$Y_t^* = \frac{1 - \lambda}{1 - \lambda L} Y_t$$

co pozwala zapisać model 12.6 w kategoriach obserwowalnego Y_t

$$C_t = a + b \frac{1 - \lambda}{1 - \lambda L} Y_t + \varepsilon_t$$

Model tej postaci jest modelem DL z geometrycznymi opóźnieniami Koycka.

Literatura: Steward (1991) str. 177-188, Green (1997) str. 782-804, Goldberger (1972) str. 352-356, Theil (1979) str. 265-269

12.2. Modele *ARDL*

Zajmiemy się teraz bardziej ogólną klasą modeli dynamicznych jaką są modele *ARDL* (**A**utoregressive **D**istributed **L**ags). Stanowią one szczególny przypadek modeli *ARMAX*, kiedy $C(L) = 1$.

Modele *ARDL* mają postać

$$A(L)y_t = B(L)x_t + \varepsilon_t \quad (12.7)$$

co można zapisać jako

$$y_t = \alpha_1 y_{t-1} + \dots + \alpha_p y_{t-p} + x_t \beta_0 + x_{t-1} \beta_1 + \dots + x_{t-s} \beta_s + \varepsilon_t \quad (12.8)$$

Częstym przekształceniem stosowanym w przypadku modeli autoregresyjnych jest przekształcenie, które pozwala przedstawić model *ARDL* jako model *DL* o nieskończonej ilości opóźnień i skorelowanym błędzie losowym

$$y_t = A(L)^{-1} B(L)x_t + A(L)^{-1} \varepsilon_t$$

Model y_t będzie stacjonarne jedynie wtedy, gdy wpływ błędu losowego ε_t będzie na y_{t+s} będzie zanikał dla dużych s . Z rozważań w dodatku 17.2 wynika, że będzie tak tylko wtedy, gdy wszystkie pierwiastki wielomianu operatora opóźnień $A(L)$ będą znajdować się poza kołem jednostkowym.

W najprostszym przypadku x_t zawiera jedynie jedną zmienną. W takim modelu jeśli x_t ulegnie zmianie o jednostkę to y_t zmieni się o β_0 . Jeśli jednak x_t pozostanie na nowym poziomie to długookresowy wpływ zmiany x_t na y_t wyniesie

$$A(1)^{-1} B(1) = \frac{\beta_0 + \beta_1 + \dots + \beta_s}{1 - \alpha_1 - \dots - \alpha_p} = \beta$$

Dla modeli *ARDL* mnożnik bezpośredni jest równy β_0 a mnożnik skumulowany jest równy β . Modele *ARDL* mające formę równania (12.8) można szacować za pomocą *MNK*. Szacowanie modeli *ARDL* jest typowym przypadkiem, w którym część zmiennych objaśniających używanych w *MNK* jest zmiennymi losowymi. Należy pamiętać, że w tym przypadku rozkłady statystyk testowych są znane jedynie dla dużych prób.

Estymatory *MNK* będą zgodne jedynie wtedy, kiedy ε_t jest nieskorelowane. Jeśli ε_t jest skorelowane, wtedy pojawi się problem równoczesności, ponieważ ε_t będzie skorelowane z opóźnionymi y_t , które w takim modelach pełnią rolę zmiennych objaśniających. W modelach *ARDL* autokorelacja ε_t może więc spowodować asymptotyczne obciążenie estymatora b_{MNK} .

Literatura: Steward (1991) str. 188-192, Green (1997) str. 804-806.

12.3. Procesy *ARIMA*

Procesy *ARIMA* (p, d, q) (**A**utoregressive **I**ntegrated **M**oving **A**verage) są to procesy, które można sprowadzić za pomocą różnic rzędu d do modelu *ARMA* (p, q). Oznacza to, że

$$A(L) \Delta^d x_t = \delta + C(L) \varepsilon_t \quad (12.9)$$

i wszystkie pierwiastki tego wielomianu opóźnień $A(L)$ leżą poza kołem jednostkowym. Sposób ustalania rzędu integracji procesu d omówimy w podrozdziale 12.4. W tej części zajmiemy się przypadkiem kiedy d jest znane. W tej sytuacji za zmienną objaśnianą możemy przyjąć $y_t = \Delta^d x_t$ i proces (12.9) zapisać jako

$$A(L) y_t = \delta + C(L) \varepsilon_t$$

czyli jako proces *ARIMA* ($p, 0, q$), który oznaczamy jako *ARMA* (p, q) (**A**utoregressive **M**oving **A**verage). Parametry p i q procesu są rzędami wielomianów operatora opóźnień $A(L)$ i $C(L)$. Proces *ARMA* (p, q) można także zapisać następująco

$$y_t + \alpha_1 y_{t-1} + \dots + \alpha_p y_{t-p} = \delta + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} \quad (12.10)$$

Jeśli wszystkie pierwiastki wielomianów $A(L)$ i $C(L)$ są znajdują się poza kołem jednostkowym, to możliwe jest przekształcenie procesu $ARMA(p, q)$ w proces $AR(\infty)$ (proces autoregresyjny o nieskończonej ilości opóźnień) za pomocą przekształcenia

$$C^{-1}(L)A(L)y_t = \delta + \varepsilon_t$$

lub też na proces $MA(\infty)$ (proces średniej ruchomej o nieskończonej ilości opóźnień) za pomocą przekształcenia

$$y_t = A^{-1}(1)\delta + A^{-1}(L)C(L)\varepsilon_t$$

W przypadku, kiedy w modelu nie występują elementy związane ze średnią ruchomą ($q = 0$) model sprowadza się do modelu autoregresyjnego $AR(p)$, który można wyestymować za pomocą standardowego *MNK*. W przypadku jednak, kiedy w modelu występują elementy związane z średnią ruchomą ($q \neq 0$), zastosowanie *MNK* prowadzi do uzyskania estymatorów asymptotycznie obciążonych z powodu występowania problemu równoczesności. W tym przypadku należy zastosować estymatory *MNW*. Do testowania hipotez prostych i złożonych dla tak wyestymowanych modeli *ARMA* można zastosować testy *LM*, *LR* lub *W*. Najważniejszym problemem, który musimy rozwiązać przy estymacji procesu *ARMA* jest ustalenie wielkości parametrów p i q . Można do tego celu zastosować kilka alternatywnych procedur. Jedną możliwością jest zastosowanie znalezienie takich p i q , dla których wartość kryteriów informacyjnych jest najniższa. Drugą możliwością jest zastosowanie do ustalenia p i q metody od ogólnego do szczegółowego w sposób analogiczny do opisanego w podrozdziale dotyczącym modeli *DL*. Ostatnim sposobem, identyfikacji p i q jest metodologia zaproponowana przez Boxa-Jenkinsa.

• Funkcje *ACF* i *PACF*

Metodologia Boxa-Jenkinsa opiera się na obserwacji, że parametry p i q procesy można wywnioskować na podstawie analizy zachowania dwóch funkcji: funkcji autokorelacji procesu *ACF* (Autocorrelation Function) i funkcji autokorelacji cząstkowej *PACF* (Partial Autocorrelation Function). Funkcja

ACF zdefiniowana jest na zbiorze liczb naturalnych. Wartość funkcji ACF dla pewnego k , oznaczane jako ρ_k , jest równa

$$\rho_k = \frac{\text{Cov}(y_t, y_{t-k})}{\text{Var}(y_t)}$$

Funkcja $PACF$ jest także zdefiniowana na zbiorze liczb naturalnych. Wartość α_{kk} tej funkcji jest równa wyestymowanemu współczynnikowi α_k w modelu $AR(k)$

$$y_t = \delta + \alpha_1 y_{t-1} + \dots + \alpha_k y_{t-k} + \varepsilon_t$$

Metodologia Boxa-Jenkinsa opiera się na obserwacji, że kształty funkcji ACF i $PACF$ zależą od wielkości parametrów p i q . Rozważmy najpierw proces $AR(1)$

$$y_t = \delta + \alpha_1 y_{t-1} + \varepsilon_t$$

Proces ten można zapisać za pomocą wielomianu operatora opóźnień

$$(1 - \alpha_1 L) y_t = \delta + \varepsilon_t$$

Korzystając ze wzoru na odwrotność wielomianu operatora opóźnień uzyskujemy

$$y_t = \frac{\delta}{1 - \alpha_1} + \sum_{i=0}^{\infty} \alpha_1^i \varepsilon_{t-i}$$

Zakładamy, ε_t są nieskorelowane. Wynika z tego, że

$$E(y_t) = \frac{\delta}{1 - \alpha_1} + \sum_{i=0}^{\infty} \alpha_1^i E(\varepsilon_{t-i}) = \frac{\delta}{1 - \alpha_1} \quad (12.11)$$

$$\text{Var}(y_t) = \text{Var}\left(\sum_{i=0}^{\infty} \alpha_1^i \varepsilon_{t-i}\right) = \sigma^2 \sum_{i=0}^{\infty} \alpha_1^{2i} = \frac{\sigma^2}{1 - \alpha_1^2}$$

Z drugiej strony

$$\begin{aligned}\text{Cov}(y_t, y_{t-k}) &= \text{E} \left[\left(\sum_{i=0}^{\infty} \alpha_1^i \varepsilon_{t-i} \right) \left(\sum_{i=0}^{\infty} \alpha_1^i \varepsilon_{t-i-k} \right) \right] \\ &= \alpha_1^k \sum_{i=0}^{\infty} \alpha_1^{2i} \sigma^2 = \alpha_1^k \frac{\sigma^2}{1 - \alpha_1^2}\end{aligned}\tag{12.12}$$

Przy czym skorzystaliśmy z niskokorelowania ε_t . Wynika z tego, że dla procesu $AR(1)$ teoretyczna wartość funkcji ACF jest równa

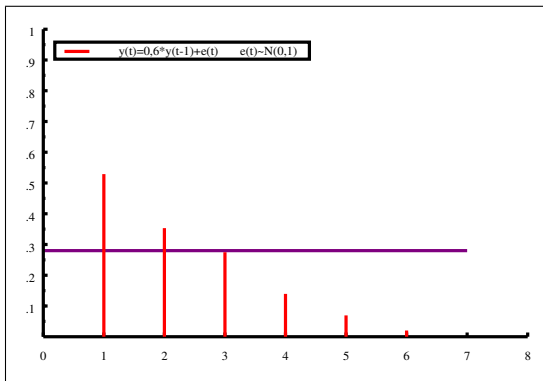
$$\rho_k = \alpha_1^k$$

przy czym łatwo zauważyć, że proces ten ma skończoną wariancję jedynie dla $|\alpha_1| < 1$. Wynika z tego, że ACF dla tego procesu maleje wykładniczo. Przykładowy przebieg funkcji ACF dla procesu $AR(1)$ przedstawiony jest na rysunku 1. Na wykresie linią przerywaną zaznaczono wartość krytyczną dla hipotezy zerowej, że $H : \rho_k = 0$.

Przebieg funkcji $PACF$ wywnioskować można z własności estymatorów. α_{11} jest estymatorem parametru przy y_{t-1} . Dla procesu $AR(1)$ jest to jedyny parametr przy opóźnionych y_t , który jest różny od zera. α_{ii} dla $i > 1$ są więc estymatorami parametrów równych zeru. Oznacza to, że jedynie α_{11} powinien być istotnie różny od zera. Przykładowy przebieg funkcji $PACF$ dla procesu $AR(1)$ zilustrowany jest na rysunku 2. Linią przerywaną zaznaczono na nim wartość krytyczną dla hipotezy, że $\alpha_{kk} = 0$.

Rozpatrzmy teraz przypadek procesu $MA(1)$ danego równaniem

$$y_t = \delta + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$



Rysunek 1: ACF dla procesu $AR(1)$

Dla takiego procesu teoretyczne wartości funkcji ACF obliczamy w następujący sposób

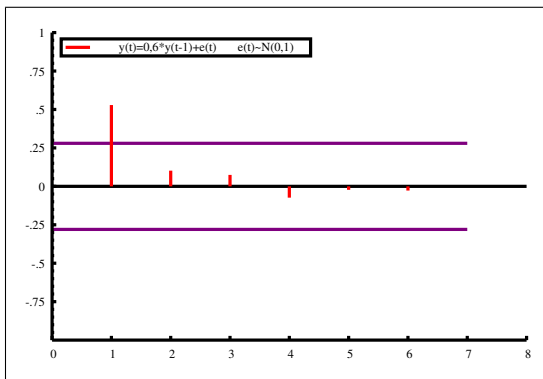
$$E(y_t) = \delta$$

$$\text{Var}(y_t) = \text{Var}(\varepsilon_t) + \theta_1^2 \text{Var}(\varepsilon_{t-1}) = (1 + \theta_1^2) \sigma^2$$

$$\text{Cov}(y_t, y_{t-1}) = E(\varepsilon_t + \theta_1 \varepsilon_{t-1})(\varepsilon_{t-1} + \theta_1 \varepsilon_{t-2})$$

$$= E(\varepsilon_t \varepsilon_{t-1}) + \theta_1 E(\varepsilon_t \varepsilon_{t-2}) + \theta_1 E(\varepsilon_{t-1}^2) + \theta_1^2 E(\varepsilon_{t-1} \varepsilon_{t-2}) = \theta_1 \sigma^2$$

Dla $i > 1$ zmienne losowe y_t i y_{t-k} są nieskorelowane ponieważ ε_t i ε_{t-1} są nieskorelowane z ε_{t-k} i ε_{t-k-1} dla $k > 1$. Wynika z tego, że teoretyczne wartości funkcji ACF dla procesu $MA(1)$



Rysunek 2: $PACF$ dla procesu $AR(1)$

kształtują się następująco

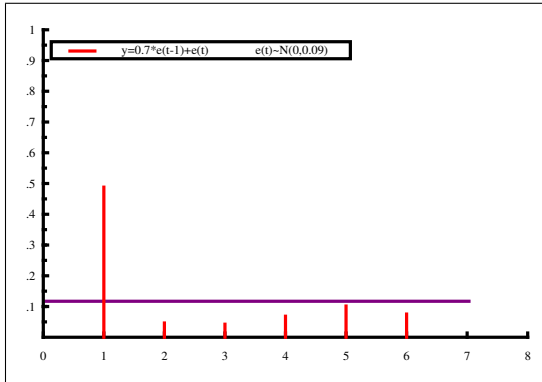
$$\rho_k = \begin{cases} \frac{\theta_1}{1+\theta_1^2} & \text{dla } k = 1 \\ 0 & \text{dla } k > 1 \end{cases}$$

Przykładowy wykres funkcji ACF dla procesu $MA(1)$ przedstawiony jest na rysunku 3.

Teoretyczne wartości funkcji $PACF$ można znaleźć przekształcając proces $MA(1)$ do postaci procesu $AR(\infty)$

$$y_t = \delta + (1 + \theta_1 L) \varepsilon_t$$

$$y_t (1 + \theta_1 L)^{-1} = (1 + \theta_1 L)^{-1} \delta + \varepsilon_t$$

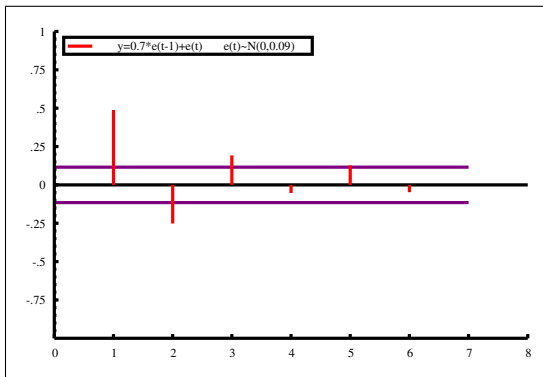


Rysunek 3: ACF dla procesu $MA(1)$

$$y_t = \frac{\delta}{1 + \theta_1} + \sum_{i=1}^{\infty} (-\theta_1)^i y_{t-i} + \varepsilon_t$$

Wynika z tego, że przebieg funkcji $PACF$ będzie w przybliżeniu dany przez $\alpha_{kk} \approx (-\theta)^k$. Wykres $PACF$ dla procesu $MA(1)$ powinien więc wykładniczo maleć. Mówimy tu jedynie o przybliżeniach, ponieważ oszacowania procesu $MA(1)$ na podstawie procesu $AR(p)$ będą asymptotycznie obciążone ze względu na występowanie problemu zmiennych pominiętych. Przykładowy wykres funkcji $PACF$ dla procesu $MA(1)$ przedstawiony jest na rysunku 4.

Nasze rozważania można w prosty sposób uogólnić na procesy $AR(p)$ i $MA(q)$. W przypadku



Rysunek 4: $PACF$ dla procesu $MA(1)$

procesów $AR(p)$

$$y_t = \delta + \alpha_1 y_{t-1} + \dots + \alpha_p y_{t-p} + \varepsilon_t$$

z rozważań z podrozdziału na temat wielomianów operatora opóźnień wynika, że

$$y_t = \frac{\delta}{1 - \alpha_1 - \dots - \alpha_p} + \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}$$

Wartość oczekiwana y_t wynosi

$$E(y_t) = \frac{\delta}{1 - \alpha_1 - \dots - \alpha_p}$$

zaś jego wariancja

$$\text{Var}(y_t) = \text{Var}\left(\sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}\right) = \sigma^2 \sum_{i=0}^{\infty} \psi_i^2 = \gamma_0$$

zaś kowariancje

$$\begin{aligned} \text{Cov}(y_t, y_{t-k}) &= \text{E}\left[\left(\sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}\right)\left(\sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i-k}\right)\right] \\ &= \sigma^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+k} = \gamma_k \end{aligned}$$

Ponieważ, jak ustaliliśmy w rozdziale (17.2) współczynniki ψ_i wykładniczo maleją więc γ_k będzie także wykładniczo maleć a tym samym funkcja ACF postaci

$$\rho_k = \frac{\gamma_k}{\gamma_0}$$

będzie także wykładniczo maleć. Wartości funkcji $PACF$ są różne 0 jedynie dla p pierwszych α_{kk} . W przypadku procesu $MA(q)$ wartości ρ_k funkcji ACF są istotnie różne od zera jedynie dla pierwszych q opóźnień a funkcja $PACF$ wykładniczo maleje. W przypadku, kiedy mamy do czynienia z procesem $ARMA(p, q)$ dla $p \neq 0$ oraz $q \neq 0$, funkcja ACF i funkcja $PACF$ po początkowych fluktuacjach zaczynają zbiegać do zera. Box i Jenkins proponowali w tym przypadku dopasowywanie procesów o coraz wyższych wielkościach parametrów p i q aż do momentu, kiedy reszty z tak dopasowanego procesu nie wykazują autokorelacji¹. Sugeruje się, że wystarczy do tego celu dopasować model, dla którego $p + q \leq 3$. Testem, który stosuje się do testowania hipotezy

¹Jednak postępując w ten sposób przechodzimy od modelu szczegółowego do ogólnego.

zerowej, że błędy losowe $\varepsilon_t, \dots, \varepsilon_{t-K}$ w takim modelu są nieskorelowane jest statystyka Ljung-Boxa

$$Q^* = n(n+2) \sum_{k=1}^K \frac{\hat{\rho}_k^2}{n-k} \xrightarrow{D} \chi_{K-p-q}^2$$

gdzie $\hat{\rho}_k$ są zdefiniowane jako korelacje reszt e_t z wyestymowanego modelu $ARMA(p, q)$

$$\hat{\rho}_k = \frac{\sum_{t=k+1}^n e_t e_{t-k}}{\sum_{t=k+1}^n e_t^2}$$

• Równania Yule-Walkera

Zauważ w modelu $AR(q)$, dla $k=0$ uzyskujemy

$$\begin{aligned} \gamma_0 &= E(y_t, y_t) = E[(a_1 y_{t-1} + \dots + a_p y_{t-p} + \varepsilon_t) y_t] \\ &= \text{Cov}(\varepsilon_t, y_t) + \sum_{i=1}^p a_i E(y_{t-i} y_t) = \sigma^2 + \sum_{i=1}^p a_i \gamma_i \end{aligned} \quad (12.13)$$

Dla $k > 0$

$$\begin{aligned} \gamma_k &= \text{Cov}(y_t, y_{t-k}) = E[(a_1 y_{t-1} + \dots + a_p y_{t-p} + \varepsilon_t) y_{t-k}], \\ &= \sum_{i=1}^p a_i E(y_{t-i} y_{t-k}) = \sum_{i=1}^p a_i \gamma_{k-i} \end{aligned} \quad (12.14)$$

przy czym korzystamy z tego, że

$$\text{Cov}(y_{t-i}, y_{t-k}) = \text{Cov}(y_t, y_{t-(k-i)}) = \gamma_{k-i}$$

i tego, że

$$\text{Cov}(y_t, y_{t-k}) = \gamma_k = \text{Cov}(y_t, y_{t+k}) = \gamma_{-k}$$

Równanie (12.14) można zastosować do znalezienia γ_k dla $k > p$ jeśli znamy γ_i dla $i = 1, \dots, p$.

• Prognozowanie za pomocą modeli *ARIMA*

Modele *ARMA* (p, q) są bardzo popularne jako proste narzędzie służące do prognozowania przyszłego zachowania szeregów czasowych. Szczególną zaletą tej metodologii jest prostota z jaką takie prognozy można formułować. Oczywistym wzorem na prognozę y_t dla okresu $t + 1$ dla procesu 12.10 jest

$$\hat{y}_{t+1} = \hat{\delta} - \hat{\alpha}_1 y_t - \hat{\alpha}_2 y_{t-1} \dots - \hat{\alpha}_p y_{t-p+1} + e_t + \hat{\theta}_1 e_{t-1} + \dots + \hat{\theta}_q e_{t-q}$$

Po to by sformułować prognozę, y_t na okres $t + 2$ możemy posłużyć się w analogicznym wzorze sformułowaną już prognozą dla \hat{y}_{t+1}

$$\hat{y}_{t+2} = \hat{\delta} - \hat{\alpha}_1 \hat{y}_{t+1} - \hat{\alpha}_2 y_t \dots - \hat{\alpha}_p y_{t-p+2} + \hat{\theta}_1 e_t + \dots + \hat{\theta}_q e_{t-q+1}$$

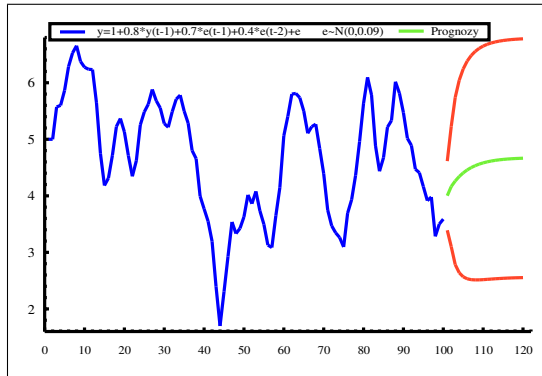
Sposób ten zastosować do znalezienia prognozy \hat{y}_{t+s} podstawiając za $y_{t+1}, \dots, y_{t-p+s}$ rekurencyjnie policzone prognozy $\hat{y}_{t+1}, \dots, \hat{y}_{t-p+s}$. Ponieważ wpływ bardziej oddalonych błędów losowych na y_t jest coraz mniejszy, więc prognozy wygenerowane dla procesu *ARMA* zbiegają, dla bardziej oddalonych w czasie okresów, do średniej tego procesu równej $\frac{\delta}{1-\theta_1-\dots-\theta_k}$. Przykładowy wykres prognoz wygenerowanych dla procesu *ARMA* (1, 1) pokazany jest na rysunku 5. Na rysunku tym zaznaczono też wartości $\hat{y}_{t+s} \pm 2\sqrt{\text{Var}(\hat{y}_{t+s})}$. Zauważmy, że podobnie jak prognozy, tak i wariancja błędu prognozy dąży z upływem czasu do stałej.

Literatura: Steward (1991) str. 209-245, Green (1997) str. 823-841.

• Estymacja procesów *ARIMA*(p, d, q)

• Estymator *MNK*

Estymację procesu *AR* (p) można przeprowadzić przy użyciu *MNK* pod warunkiem, że błędy losowe są nieskorelowane i homoskedystyczne. W takim przypadku można przyjmujemy za zmienną zależną y_t a zmiennymi niezależnymi są zmienne y_{t-1}, \dots, y_{t-p} . Konieczne jest pominięcie pierwszych p okresów, ponieważ jeśli y_1 jest pierwszą obserwacją, to brak nam informacji na temat



Rysunek 5: Prognozy z modelu $ARMA(1, 2)$

y_0, \dots, y_{-p+1} , których poziomy musimy znać jako, że są one poziomami zmiennych niezależnych dla okresów $t = 1, \dots, p$. Oznacza to, że dodawanie opóźnień w modelu prowadzi do utraty stopni swobody nie tylko poprzez konieczność szacowania większej ilości parametrów ale i z powodu ograniczania wielkości próby.

Ponieważ zmienne niezależne w takim modelu są losowe więc możemy posługiwać się jedynie wynikami asymptotycznymi. Stosowanie statystyk t i F może być jednak w małych próbach uzasadnione o ile pokazano, że są przyzwoitymi one aproksymacjami prawdziwych rozkładów małopróbkowych.

Przy estymacji MNK modelu $AR(p)$ trzeba szczególnie zwracać uwagę na problem autokorelacji. Jak pokazaliśmy w przykładzie (8.1), w tym przypadku autokorelacja może doprowadzić do wystąpienia problemu równoczesności, który powoduje, że estymator, nawet w dużej próbie będzie

obciążony.

• Estymator Yule-Walkera

Równania (12.14) można zapisać w postaci macierzowej jako

$$\begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_k \end{bmatrix} = \begin{bmatrix} \gamma_0 & \cdots & \gamma_{p-1} \\ \vdots & \ddots & \vdots \\ \gamma_{p+1} & \cdots & \gamma_0 \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix}$$

Oczywistym kandydatem na estymator będzie

$$\begin{bmatrix} \hat{a}_1 \\ \vdots \\ \hat{a}_p \end{bmatrix} = \begin{bmatrix} \hat{\gamma}_0 & \cdots & \hat{\gamma}_{p-1} \\ \vdots & \ddots & \vdots \\ \hat{\gamma}_{p+1} & \cdots & \hat{\gamma}_0 \end{bmatrix}^{-1} \begin{bmatrix} \hat{\gamma}_1 \\ \vdots \\ \hat{\gamma}_k \end{bmatrix}$$

Gdzie korelacje empiryczne $\hat{\gamma}_k$ są wyznaczone dla $k = 0, \dots, p$ na jako momenty z próby $\hat{\gamma}_k$

$$\bar{y} = T^{-1} \sum_{t=1}^T y_t$$
$$\hat{\gamma}_k = T^{-1} \sum_{t=k+1}^T (y_t - \bar{y})(y_{t-k} - \bar{y})$$

. Estymator σ^2 można znaleźć z równania (12.13) przyjmując, że

$$\hat{\sigma}^2 = \hat{\gamma}_0 - \sum_{i=1}^p \hat{a}_i \hat{\gamma}_i$$

Estymator Yule-Walkera parametrów a_1, \dots, a_k jest równoważny estymatorowi *MNK* jeśli do policzenia momentów z próby wykorzystamy jedynie obserwacje y_{p+1}, \dots, y_T .

12.4. Integracja i kointegracja

• Procesy stacjonarne i niestacjonarne

Wszystkie procesy stochastyczne, które dotąd analizowaliśmy były procesami stacjonarnymi. Definicje stacjonarności i niestacjonarności znajdują się w podrozdziale 13.5. Typowym przypadkiem skalarnych procesów stacjonarnych są procesy $ARMA(p, q)$, ponieważ jak pokazaliśmy, w ich przypadku $\text{Var}(y_t) = \text{const}$ a $\text{Cov}(y_t, y_{t-k}) = \gamma(k)$, które zależy wyłącznie od k .

Najczęściej rozpatrywanym przykładem procesu niestacjonarnego proces

$$y_t = y_{t-1} + \varepsilon_t \quad (12.15)$$

Jeśli podstawimy w tym wzorze za $y_{t-1} = y_{t-2} + \varepsilon_{t-1}$, to uzyskamy $y_t = y_{t-2} + \varepsilon_{t-1} + \varepsilon_t$. Postępując rekurencyjnie w ten sposób uzyskamy, że

$$y_t = y_0 + \sum_{s=1}^t \varepsilon_s$$

Jeśli założymy, że $y_0 = 0$, to

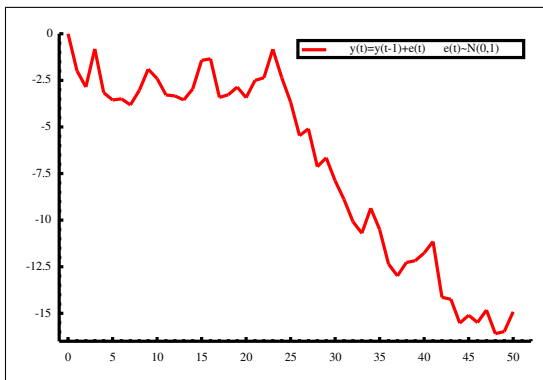
$$E(y_t) = 0$$

$$\text{Var}(y_t) = \sum_{s=1}^t \text{Var}(\varepsilon_s) = t\sigma^2$$

$$\text{Cov}(y_t, y_{t+h}) = \sum_{s=1}^t \text{Var}(\varepsilon_s) + 2 \sum_{i \neq j} \text{Cov}(\varepsilon_i, \varepsilon_j) = t\sigma^2$$

Zauważmy, że w tym przypadku zarówno wariancja jak i kowariancja procesu nie są stałe w czasie co oznacza, że proces ten rzeczywiście nie jest procesem stacjonarnym. Przykładowy wykres procesu (12.15) znajduje się na rysunku 6. Innym przykładem procesów niestacjonarnych są procesy

ARIMA(p, d, q) dla $d > 0$.



Rysunek 6: Zmienna $x_t \sim I(1)$

Ważkość pytania, czy proces jest stacjonarny wynika z wpływu stacjonarności na własności estymatorów. Rozważmy na przykład założenie 6.3. Dla procesów stacjonarnych można, przy użyciu Centralnego Twierdzenia Granicznego, pokazać, że macierz

$$T^{-1} \mathbf{X}' \mathbf{X} \xrightarrow{P} \text{Var}(\mathbf{x}_t) - \text{E}(\mathbf{x}_t) \text{E}(\mathbf{x}_t') = \mathbf{Q} = \text{const}$$

Dla procesów niestacjonarnych nie można uzyskać podobnego wyniku, ponieważ $\text{Var}(\mathbf{x}_t)$ nie jest stałą. Przykładowo dla procesu (12.15) macierz $\text{plim}(\text{Var}(\mathbf{x}_t)) = \infty$. W rezultacie wnioski z podrozdziału 6.3 nie mają zastosowania dla procesów niestacjonarnych.

• Rząd integracji zmiennych

Szczególnym rodzajem procesów niestacjonarnych są procesy, które można sprowadzić do stacjonarności z pomocą różnicowania. Najprostszym przykładem takiego procesu jest proces (12.15). Zauważmy, że dla tego procesu

$$\Delta y_t = \varepsilon_t$$

jest rzeczywiście stacjonarna. Proces, który można doprowadzić do stacjonarności za pomocą d -tych różnic nazywamy zintegrowanym rzędu d i oznaczamy jako $y_t \sim I(d)$. Przykładowo takimi procesami są procesy $ARIMA(p, d, q)$. Dla procesów takich jak procesy $ARIMA(p, d, q)$, które można zapisać za pomocą wielomianu operatora opóźnień jako

$$A(L) y_t = \delta + \varepsilon_t \quad (12.16)$$

wartość oczekiwana jest równa

$$E(y_t) = E[A^{-1}(L)(\delta + \varepsilon_t)] = A^{-1}(1)(\delta + \varepsilon_t)$$

a wariancja

$$\text{Var}(y_t) = E[A^{-1}(L)\varepsilon_t^2] = A^{-2}(1)\sigma^2$$

z rozważań na temat własności wielomianów operatora opóźnień wiemy, że $A^{-1}(1)$ jest skończone jedynie dla przypadku, kiedy pierwiastki wielomianu $A(L)$ leżą poza kołem jednostkowym. Zauważmy teraz, że warunek ten jest także warunkiem koniecznym, by $E(y_t)$ i $A^{-2}(1)\sigma^2$ były skończone a proces y_t stacjonarny.

Integracja rzędu d dla procesów typu (12.16) oznacza, że można je przekształcić do

$$\Delta^d y_t = \Delta^d A^{-1}(L)\varepsilon_t = (1-L)^d A^{-1}(L)\varepsilon_t = C_1(L)\varepsilon_t$$

gdzie $C_1(L)\varepsilon_t$ jest stacjonarne. Wynika z tego, że $A(L) = (1-L)^d C_1(L)$, co oznacza, że wielomian $A(L)$ ma d pierwiastków jednostkowych.

W świetle tego, co dotychczas powiedzieliśmy bardzo istotnym problem jest problem znalezienia stopnia integracji procesu. Ponieważ ustalenie rzędu integracji równoważne jest ustaleniu ilości pierwiastków równych zero dla wielomianu $A(L)$ testy te nazywamy testami pierwiastka jednostkowego (*unit root tests*).

• Test Dickey-Fullera

Najpopularniejszym obecnie testem na *istnienie* pierwiastka jednostkowego jest test DF (Dickey-Fuller). Hipotezą zerową jest, że proces $y_t \sim I(1)$. W najprostszych przypadkach hipoteza ta jest równoważna hipotezie, że w równaniu

$$y_t = \alpha y_{t-1} + \varepsilon_t \quad (12.17)$$

$\alpha = 1$. Okazuje się, że prościej jest estymować przekształcone równanie

$$\Delta y_t = (\alpha - 1) y_{t-1} + \varepsilon_t = \rho y_{t-1} + \varepsilon_t \quad (12.18)$$

ponieważ w tym przypadku hipoteza zerowa o istnieniu pierwiastka jednostkowego sprowadza się do $H_0 : \rho = 0$. Okazuje się, że rozkład tego testu nie jest rozkładem t -studenta. Specjalne tablice wartości krytycznych można znaleźć między innymi w książce Charemzy i Deadmana (1997). Dla nieco ogólniejszego przypadku, dla którego ε_t może być skorelowane zamiast równania 12.18 estymujemy równanie

$$\Delta y_t = \rho y_{t-1} + \sum_{i=1}^k \theta_i \Delta y_{t-i} + \varepsilon_t \quad (12.19)$$

przy czym k jest tak dobrane, że reszty z wyestymowanego równania były nieskorelowane. Tak uogólniony test DF nazywamy rozszerzonym testem Dickey-Fullera ADF (Augmented Dickey-Fuller). Rozkład tego testu jest dokładnie taki sam jak testu DF .

Jeśli podejrzewamy, że rząd d integracji procesu jest większy od 1 możemy posłużyć się procedurą Dickey-Pantuli. Najpierw dobieramy najwyższy rząd integracji s jaki według nas może mieć

analizowany proces. Sugeruje się przy tym, że w ekonomii nie spotyka się szeregów o rzędzie integracji wyższym niż 3. Następnie kolejno testujemy hipotezy, że $\Delta^s x_t, \dots, \Delta x_t$ zawierają pierwiastek jednostkowy. Za rząd integracji procesu przyjmujemy najwyższe i , dla którego hipoteza ta została *nie została odrzucona*. Ostatnia uwaga dotyczy roli trendów deterministycznych przy testowaniu hipotezy o istnieniu pierwiastka jednostkowego. Zauważmy, że jeśli

$$\Delta y_t = \delta + \rho y_{t-1} + \varepsilon_t$$

to w przypadku prawdziwości $H_0 : \rho = 0$, mamy $y_t = \alpha + \delta t + \sum_{s=1}^t \varepsilon_t$, a jeśli hipoteza zerowa jest fałszywa, to $y_t = \frac{\delta}{1-\alpha} + \sum_{s=1}^t \alpha^s \varepsilon_t$. Kształt trendu deterministycznego dla hipotezy zerowej i alternatywnej są więc różne². Co ważniejsze, rozkłady testów DF i ADF zależą od formy trendu deterministycznego w modelu. Przed przystąpieniem do testowania istnienia pierwiastka jednostkowego należy więc zdecydować, czy i jakiej postaci trend deterministyczny występuje w analizowanym procesie. Natępnie należy wyestymować parametry tego procesu i porównać uzyskaną wartość statystyki z tablicami testu ADF policzonymi dla procesów z tak zdefiniowanym trendem.

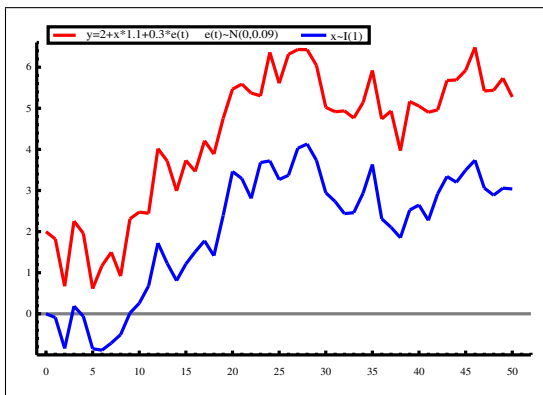
• Kointegracja

Niestacjonarność zmiennych może być powodem istotnych trudności ponieważ w przypadku modeli zawierających takie zmienne nie są spełnione założenia konieczne dla uzyskania wyników z podrozdziału 6.3. Szczególnie duże problemy może spowodować stosowanie standardowych rozkładów w przeprowadzeniu w takich modelach testów istotności. Pokazano, że dla modelu jednorównaniowego, H_0 o nieistotności zmiennej niezależnej jest asymptotycznie zawsze odrzucana jeśli tylko zmienna objaśniana i zmienna objaśniająca są $I(1)$. Nie ma przy tym znaczenia, czy istnieje rzeczywisty związek między tymi dwoma zmiennymi. Regresję z takimi fałszywie istotnymi statystykami nazywamy regresjami rzekomymi (*spurious regressions*).

Istnieją jednak ważne przypadki, dla których estymacja modeli dla zmiennych niestacjonarnych

²Jest to jedna z poważnych wad testu Dickey-Fullera.

jest możliwa a uzyskane wyniki mają głęboką interpretację. Szczególnie ważne jest tu pojęcie kointegracji. O procesie $y_t \sim I(d)$ i procesie $x_t \sim I(d)$ mówimy, że są skointegrowane rzędu b z wektorem kointegrującym $(1, -\beta)'$ jeśli $y_t - x_t\beta$ jest zintegrowane rzędu $d - b$. Oznaczamy to jako $(y_t, x_t) \sim CI(d, b)$. Najczęściej analizowanym przypadkiem jest $(y_t, x_t) \sim CI(1, 1)$, kiedy $y_t \sim I(1)$ i $x_t \sim I(1)$ a $y_t - x_t\beta \sim I(0)$. Przykładowy przebieg tak skointegrowanych szeregów przedstawia rysunek 7.



Rysunek 7: Proces $CI(1, 1)$

Przypadek $(y_t, x_t) \sim CI(1, 1)$ jest szczególnie interesujący, ponieważ udowodniono następujące twierdzenie.

Twierdzenie 12.2 (Twierdzenie Grangera) Jeśli $(y_t, x_t) \sim CI(1, 1)$ to y_t można przedstawić w

postaci następującego Mechanizmu Korekty Błędów ECM (**Error Correction Mechanism**)

$$\Delta y_t = \alpha (y_t - \mathbf{x}_t \boldsymbol{\beta}) + \sum_{i=1}^k \theta_i \Delta y_{t-i} + \sum_{i=1}^k \Delta \mathbf{x}_{t-i} \boldsymbol{\gamma}_i + \varepsilon_t \quad (12.20)$$

Mechanizm korekty błędów ma ciekawą z punktu widzenia ekonomii interpretację. Wyrażenie $y_t = \mathbf{x}_t \boldsymbol{\beta}$ interpretujemy jako równowagę długookresową, a $y_t - \mathbf{x}_t \boldsymbol{\beta}$ jako odchylenie od tej równowagi. Współczynnik α związany jest z szybkością dostosowania y_t do poziomu równowagi. Współczynniki θ_i i $\boldsymbol{\gamma}_i$ związane są z krótkookresową dynamiką procesu.

Większość zależności między zmiennymi ekonomicznymi wynikająca z teorii ekonomii ma charakter długookresowy. Mechanizm korekty błędów umożliwia rozróżnienie takich długookresowych zależności od zależności związanych z dynamiką krótkookresową.

W przypadku tego typu analizy musimy rozwiązać dwa współzależne problemy. Po pierwsze, stwierdzić, czy istnieje kointegracja między zmiennymi. Po drugie, jeśli proces jest skointegrowany, należy wyestymować parametry procesu 12.20. Poniżej omówimy najczęściej w praktyce stosowaną metodę postępowania.

• Dwustopniowa procedura Engla-Grangera

W dwustopniowej metodzie Engla-Grangera wykorzystujemy fakt, że jeśli $(y_t, \mathbf{x}_t \boldsymbol{\beta}) \sim CI(1, 1)$, to MNK zastosowany do modelu

$$y_t = \mathbf{x}_t \boldsymbol{\beta} + \varepsilon_t \quad (12.21)$$

da zbieżny estymator parametru $\boldsymbol{\beta}$. O estymatorze tym mówi się, że jest superzbieżny, ponieważ $T^{\frac{1}{2}} \mathbf{b} \xrightarrow{P} \boldsymbol{\beta}$, podczas gdy w standardowym przypadku $y_t \sim I(0)$, $\mathbf{x}_t \sim I(0)$ mamy jedynie, że $\mathbf{b} \xrightarrow{P} \boldsymbol{\beta}$. Na pierwszym etapie procedury Engla-Grangera wyliczamy estymator \mathbf{b}_{MNK} .

Możemy teraz przetestować, czy y_t i \mathbf{x}_t są skointegrowane. Jeśli y_t i \mathbf{x}_t są skointegrowane, to reszty $\hat{\varepsilon}_t = y_t - \mathbf{x}_t \mathbf{b}$ są dla dużych prób stacjonarne. Test Engla i Grangera na kointegrację sprawdza się do zastosowania testu ADF do $\hat{\varepsilon}_t$. Rozkład tego testu zależał będzie zarówno od

ilości elementów wektora \mathbf{x}_t jak i od tego, czy w \mathbf{x}_t znajduje się stała lub trend deterministyczny. Jeśli stwierdzimy, że $y_t \sim I(1)$, $\mathbf{x}_t \sim I(1)$, ale reszty $\hat{\epsilon}_t$ z modelu (12.21) są stacjonarne, to y_t i \mathbf{x}_t są skointegrowane.

Ponieważ \mathbf{b} jest superzbieżny więc można go w procesie estymacji traktować jako znaną stałą. Podstawiając \mathbf{b} do równania (12.20) uzyskujemy równanie w którym wszystkie zmienne, to jest $\Delta y_t, y_t - \mathbf{x}_t \mathbf{b}, \Delta y_{t-1}, \dots, \Delta y_{t-k}, \Delta \mathbf{x}_{t-1}, \dots, \Delta \mathbf{x}_{t-k}$ są stacjonarne. Na drugim etapie procedury Engla-Grangera estymujemy więc równanie

$$\Delta y_t = \alpha \hat{\epsilon}_t + \sum_{i=1}^k \theta_i \Delta y_{t-i} + \sum_{i=1}^k \Delta \mathbf{x}_{t-i} \gamma_i + \varepsilon_t$$

za pomocą standardowego MNK.

Przykład 12.3 *Testowanie hipotezy o parytecie siły nabywczej PPP (Purchasing Power Parity). Jedną z bardziej znanych hipotez ekonomicznych jest hipoteza, że siła kursy walut powinny być takie, by siła nabywczą określonej waluty, po przeliczeniu na inną walutę, miała taką samą siłę nabywczą za granicą jak w kraju. Jeśli hipoteza ta jest prawdziwa jest, to*

$$\frac{E_t}{I_t} = \frac{E_t^*}{I_t^*},$$

gdzie E_t jest ceną US\$ w walucie krajowej, E_t^* ceną US\$ w walucie zagranicznej a I_t, I_t^* są indeksami cen odpowiednio w kraju i zagranicą. Po zlogarytmowaniu stronami uzyskujemy

$$e_t - i_t = e_t^* - i_t^*$$

Jest prawdopodobne, że PPP obowiązuje jedynie w długim okresie. Testy pierwiastka jednostkowego wykazują, że e_t, e_t^*, i_t, i_t^* są wszystkie $I(1)$. Hipotezę o istnieniu długookresowym istnieniu PPP można przetestować badając, czy e_t, i_t, e_t^*, i_t^* są skointegrowane dla wektora kointegrującego $(1, -1, -1, 1)$.

Literatura: Steward (1991) str. 199-208, Green (1997) str. 841-859, Charemza i Deadman (1997) str. 103-144