

Rozdział 4. Jednowymiarowe modele szeregów czasowych

MODELOWANIE POLSKIEJ GOSPODARKI z R

Jednowymiarowe modele szeregów czasowych

Jednowymiarowe modele szeregów czasowych:

Klasa specyfikacji modeli ekonometrycznych, w których wnioskowanie na temat dynamiki zmiennej ekonomicznej jest przeprowadzane jedynie na podstawie przeszłych i bieżących obserwacji dla tej zmiennej.

Zastosowania:

- Analiza dynamicznych własności szeregów czasowych
- Prognozowanie

Najpopularniejsze modele jednorównaniowe:

Autoregresyjne modele średniej ruchomej (AutoRegressive Moving Average, ARMA).

Model średniej ruchomej

Proces **średniej ruchomej** (moving average) $MA(Q)$ stanowi **średnią ważoną** procesu białego szumu z bieżącego oraz Q poprzednich okresów:

$$y_t = \alpha_0 + \epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_Q \epsilon_{t-Q}, \quad \epsilon_t \sim WN(0, \sigma_\epsilon^2)$$

Inny zapis procesu $MA(Q)$ z wykorzystaniem operatora opóźnień:

$$y_t = \alpha_0 + \theta(L)\epsilon_t$$

gdzie $\theta(L) = 1 + \theta_1 L + \dots + \theta_Q L^Q$ jest wielomianem Q -tego stopnia.

Model średniej ruchomej

Momenty procesu MA(Q):

$$E(y_t) = \alpha_0$$

$$\text{Var}(y_t) = (1 + \theta_1^2 + \dots + \theta_Q^2)\sigma_\epsilon^2$$

$$\text{Cov}(y_t, y_{t-k}) = \begin{cases} 0 & \text{dla } k > Q \\ (\theta_k + \theta_{k+1}\theta_1 + \dots + \theta_Q\theta_{Q-k})\sigma_\epsilon^2 & \text{dla } k \leq Q \end{cases}$$

Uwaga: ACF wynosi zero dla opóźnień $k > Q$.

Model autoregresyjny

Dla **procesu autoregresyjnego** (autoregression) $AR(P)$ bieżąca wartość zmiennej y_t zależy od jej P przeszłych realizacji $\{y_{t-p} : p = 1, 2, \dots, P\}$:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \dots + \alpha_P y_{t-P} + \epsilon_t, \quad \epsilon_t \sim WN(0, \sigma_\epsilon^2)$$

Lub w zapisie z operatorem opóźnień:

$$\alpha(L)y_t = \alpha_0 + \epsilon_t,$$

gdzie $\alpha(L) = (1 - \alpha_1 L - \dots - \alpha_P L^P)$ jest wielomianem P -tego stopnia.

Model autoregresyjny

Proces $AR(P)$ jest **stacjonarny** tylko dla wartości parametrów α_p ($p = 1, 2, \dots, P$) dla których pierwiastki równania:

$$\alpha(z) = (1 - \lambda_1 z)(1 - \lambda_2 z) \dots (1 - \lambda_P z) = 0.$$

są co do modułu większe od 1, tj. jeżeli $|z_p| > 1$ dla $p = 1, 2, \dots, P$. Jest to równoważne warunkowi $|\lambda_p| < 1$.

Procesu $AR(1)$ postaci $y_t = \alpha_1 y_{t-1} + \epsilon_t$ jest stacjonarny gdy $|\alpha_1| < 1$.

Przykład:

- Proces $y_t = 1, 3y_{t-1} - 0, 4y_{t-2} + \epsilon_t$ jest stacjonarny ponieważ pierwiastki równania $(1 - 0, 5z)(1 - 0, 8z) = 0$ wynoszą $z_1 = 2$ oraz $z_2 = 1, 25$
- Proces $y_t = 1, 5y_{t-1} - 0, 5y_{t-2} + \epsilon_t$ jest niestacjonarny, ponieważ pierwiastki wynoszą $z_1 = 2$ oraz $z_2 = 1$.

Model autoregresyjny

Warunek stacjonarności procesu AR można także zdefiniować w kategoriach możliwości zapisania go w postaci **nieskończonej średniej ruchomej**:

$$y_t = \alpha^{-1}(L)(\alpha_0 + \epsilon_t) = \mu + \epsilon_t + \sum_{q=1}^{\infty} \theta_q \epsilon_{t-q},$$

dla której zachodzi warunek

$$\lim_{k \rightarrow \infty} \theta_k = 0$$

Oznacza to, że dynamiczny wpływ szoku ϵ na zmienną y jest tymczasowy i **wygasa w czasie**.

Powyższa definicja stacjonarności pokrywa się z **twierdzeniem Wolda** mówiącym, że każdy szereg stacjonarny można zapisać jako sumę procesu deterministycznego oraz stochastycznego w postaci nieskończonego procesu MA.

Model autoregresyjny

Momenty procesu AR(P):

$$E(y_t) = \frac{\alpha_0}{1 - \sum_{p=1}^P \alpha_p}$$

$$\text{Var}(y_t) = (1 + \sum_{q=1}^{\infty} \theta_q^2) \sigma_\epsilon^2$$

Równania Yule'a-Walkera dla wartości autokowariancji γ_k :

$$\gamma_0 = \alpha_1 \gamma_1 + \dots + \alpha_Q \gamma_Q + \sigma_\epsilon^2$$

$$\gamma_k = \alpha_1 \gamma_{k-1} + \dots + \alpha_Q \gamma_{k-Q}$$

Dla procesu AR(1):

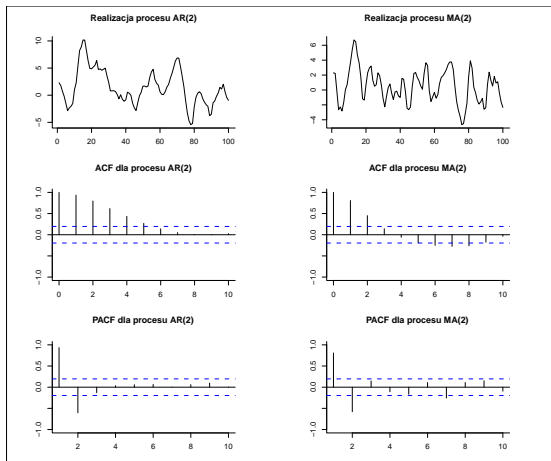
$$\gamma_k = \frac{(\alpha_1)^k}{1 - \alpha_1^2} \sigma_\epsilon^2,$$

Wartości PACF dla procesu AR(P) przyjmują wartości zerowe dla opóźnień większych od P , $\rho_{kk} = 0$ dla $k > P$.

Przykład

Proces AR: $y_t = 1,3y_{t-1} - 0,4y_{t-2} + \epsilon_t$

Proces MA: $y_t = \epsilon_t + 1,5\epsilon_{t-1} + \epsilon_{t-2}$



Model ARMA

Połączeniem modelu AR(P) oraz modelu MA(Q) jest **autoregresyjny model średniej ruchomej** ARMA(P, Q) postaci:

$$\alpha(L)y_t = \alpha_0 + \theta(L)\epsilon_t, \epsilon_t \sim WN(0, \sigma_\epsilon^2)$$

gdzie $\alpha(L) = 1 - \alpha_1L - \dots - \alpha_PL^P$ oraz $\theta(L) = 1 + \theta_1L + \dots + \theta_QL^Q$

Proces ARMA jest **stacjonarny** jeżeli pierwiastki równania $\alpha(z) = 0$ leżą poza kołem jednostkowym.

Jeżeli y_t jest procesem niestacjonarnym zintegrowanym w stopniu D , stosuje się proces ARIMA(P, D, Q) postaci:

$$\alpha(L)(1 - L)^D y_t = \alpha_0 + \theta(L)\epsilon_t.$$

Estymacja parametrów modelu AR

Dla modelu AR(1):

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \epsilon_t,$$

parametry α_0 i α_1 można oszacować **MNK**. Ze względu na **brak danych dla y_0** , dopasowanie modelu pomija błąd dla pierwszej obserwacji.

Dla $\epsilon_t \sim N(0, \sigma_\epsilon^2)$ parametry modelu AR(1) można oszacować stosując **MNW**. Niech $\mathcal{Y}_t = \{y_1, y_2, \dots, y_t\}$ - informacja dostępną do momentu t oraz $\Phi = (\alpha_0, \alpha_1, \sigma_\epsilon^2)$ - parametry modelu. **Warunkowa funkcja gęstości** dla y_t wynosi:

$$p(y_t | \Phi, \mathcal{Y}_{t-1}) = \frac{1}{\sqrt{2\pi\sigma_\epsilon^2}} \exp\left(-\frac{(y_t - \alpha_0 - \alpha_1 y_{t-1})^2}{2\sigma_\epsilon^2}\right),$$

zaś funkcja gęstości dla **wszystkich obserwacji** jest równa:

$$p(y_1, y_2, \dots, y_T | \Phi) = p(y_T | \Phi, \mathcal{Y}_{T-1}) \times p(y_{T-1} | \Phi, \mathcal{Y}_{T-2}) \times \dots \times p(y_2 | \Phi, \mathcal{Y}_1) \times p(y_1 | \Phi).$$

Znamy pierwsze $T - 1$ czynników powyższego iloczynu, oraz nie znamy wartości $p(y_1 | \Phi)$.

Estymacja parametrów modelu AR

Jeżeli założymy, że $p(y_1|\Phi)$ jest dane przez bezwarunkowy rozkład dla procesu AR(1), $y_1 \sim N(\frac{\alpha_0}{1-\alpha_1}, \frac{\sigma_\epsilon^2}{1-\alpha_1^2})$, czyli:

$$p(y_1|\Phi) = \frac{1}{\sqrt{2\pi \frac{\sigma_\epsilon^2}{1-\alpha_1^2}}} \exp\left(-\frac{(y_1 - \frac{\alpha_0}{1-\alpha_1})^2}{2 \frac{\sigma_\epsilon^2}{1-\alpha_1^2}}\right).$$

to możemy wyznaczyć **dokładną wartość funkcji wiarygodności** (exact likelihood):

$$\mathcal{L}(\Phi; y_1, \dots, y_T) = p(y_1, y_2, \dots, y_T|\Phi).$$

Maksimum powyższego wyrażenia znajduje się w sposób numeryczny (brak analitycznego rozwiązania). Z tego względu, stosuje się także **warunkową funkcję wiarygodności** (conditional likelihood):

$$\mathcal{L}(\Phi; y_2, \dots, y_T|y_1) = p(y_T|\Phi, \mathcal{Y}_{T-1}) \times p(y_{T-1}|\Phi, \mathcal{Y}_{T-2}) \times \dots \times p(y_2|\Phi, \mathcal{Y}_1).$$

W tym przypadku oszacowania α_0 oraz α_1 są takie same jak dla MNK.

Estymacja parametrów modelu MA:

Metoda warunkowej sumy kwadratów reszt (CSS)

Dla modelu MA(1) $y_t = \epsilon_t + \theta\epsilon_{t-1}$ estymacja MNK nie jest możliwa.

W celu ilustracji metody CSS zapiszmy składnik losowy procesu MA(1) jako:

$$\epsilon_t = y_t - \theta y_{t-1} + \theta^2 y_{t-2} - \dots + (-1)^{t-1} \theta^{t-1} y_1 + (-1)^t \theta^t \epsilon_0 = h(y_1, y_2, \dots, y_t, \theta, \epsilon_0),$$

W metodzie CSS przyjmuje się, że $\epsilon_0 = 0$, a tym samym reszty są równe:

$$e_t = h(y_1, y_2, \dots, y_t, \hat{\theta}, 0),$$

gdzie $\hat{\theta}$ jest oszacowaniem θ .

Suma kwadratów reszt jest nieliniową funkcją obserwacji y_t oraz parametru $\hat{\theta}$:

$$\sum_{t=1}^T e_t^2 = S(y_1, \dots, y_T; \hat{\theta}).$$

Estymacja parametrów modelu MA: Metoda warunkowej sumy kwadratów reszt (CSS)

Estymator CSS jest wartością parametru $\hat{\theta}$ który minimalizuje wartość wyrażenia $S(y_1, \dots, y_T; \hat{\theta})$.

Dla modeli ARMA(P, Q) metoda CSS zakłada, że $\epsilon_s = 0, s \leq P$, zaś minimalizowana jest suma kwadratów reszt dla $T - P$ obserwacji. Przy założeniu o normalności ϵ_t oszacowania uzyskane metodą CSS pokrywają się z oszacowaniami uzyskanymi na postawie warunkowej MNW.

W praktyce parametry modeli ARMA najczęściej szacuje się poprzez maksymalizację [dokładnej funkcji wiarygodności](#)

Ponieważ oszacowania uzyskuje się [metodami numerycznymi](#), należy ustalić [początkowe wartości parametrów](#). Sugeruje się rozpoczęcie iteracji z kilku punktów startowych lub dla wartości parametrów uzyskanych metodą CSS.

Przykład estymacji modelu ARMA

Oszacowanie modelu ARMA(3,3) dla miesięcznej inflacji CPI w Polsce (próba 1999:1-2010:10)

	ar1	ar2	ar3	ma1	ma2	ma3	cons.
CSS-ML	0.029	0.877	0.009	0.463	-0.626	-0.262	0.336
CSS	0.151	0.882	-0.130	0.374	-0.848	-0.164	0.202
ML	0.067	-0.004	0.771	0.355	0.320	-0.666	0.330

Wybór specyfikacji modelu ARMA

Metoda Boxa-Jenkinsa (1970):

- Arbitralne ustalenie wartości P i Q na podstawie wykresu y_t , funkcji ACF i PACF
- Estymacja parametrów modelu ARMA(P, Q).
- Weryfikacja modelu ARMA(P, Q) składająca się z dwóch części:
 - a. Dla modelu ARMA(P^*, Q^*), gdzie $P^* \geq P$ oraz $Q^* \geq Q$, weryfikuje się hipotezę, że dodatkowe opóźnienia nie są istotne
 - b. Sprawdza się czy występuje autokorelacja reszt

Wybór specyfikacji modelu ARMA

Kryteria informacyjne: Kryteria informacyjne faworyzują specyfikacje o wysokiej wartości logarytmu funkcji wiarygodności $\ell = \ln \mathcal{L}$ i niskiej liczbie szacowanych parametrów $K = P + Q + 1$.

Trzy najczęściej stosowane kryteria to kryterium Akaike'a (AIC), Schwarza (BIC) oraz Hannana-Quinna (HQIC):

$$AIC = -2 \frac{\ell}{T} + 2 \frac{K}{T}$$

$$BIC = -2 \frac{\ell}{T} + \frac{K}{T} \ln T$$

$$HQIC = -2 \frac{\ell}{T} + 2 \frac{K}{T} \ln(\ln T),$$

gdzie $K(SIC) \leq K(HQIC) \leq K(AIC)$

Przykład

Kryterium AIC

	ma0	ma1	ma2	ma3	ma4
ar0	64.725	20.159	6.029	5.701	4.648
ar1	-4.386	-6.265	-10.667	-9.248	-9.139
ar2	-4.286	-10.420	-8.918	-9.343	-6.445
ar3	-3.786	-9.372	-7.429	-7.344	-10.836
ar4	-7.467	-9.261	-8.128	-8.393	-7.173

Kryterium BIC

	ma0	ma1	ma2	ma3	ma4
ar0	70.622	29.005	17.824	20.444	22.340
ar1	4.438	5.500	4.040	8.401	11.451
ar2	7.451	4.252	8.687	11.198	17.030
ar3	10.849	8.191	13.060	16.073	15.509
ar4	10.052	11.178	15.231	17.886	22.026

Test ilorazu wiarygodności

Zestaw hipotez:

$$H_0 : \forall_{P < p \leq P^*} \alpha_p = 0 \wedge \forall_{Q < q \leq Q^*} \theta_q = 0$$

$$H_1 : \exists_{P < p \leq P^*} \alpha_p \neq 0 \vee \exists_{Q < q \leq Q^*} \theta_q \neq 0$$

Liczba tych restrykcji: $m = P^* - P + Q^* - Q$.

Statystyka **testu ilorazu wiarygodności** (likelihood ratio test).

$$LR = -2(\ell_r - \ell_u)$$

przy prawdziwości H_0 ma asymptotyczny rozkład χ^2 o m stopniach swobody.

Przykład: model ARMA(3,4) vs model ARMA(1,2):

Wyniki testu ilorazu funkcji wiarygodności:

LR = 8.168, prob = 0.0856

Test portmanteau na autokorelację reszt

Hipoteza zerowa:

$$H_0 : \forall_{1 \leq j \leq J} \rho_{e,j} = 0$$

$$H_1 : \exists_{1 \leq j \leq J} \rho_{e,j} \neq 0$$

Przy prawdziwości H_0 statystyka **Boxa-Pierca**:

$$BP = T \sum_{j=1}^J \hat{\rho}_{e,j}^2,$$

ma asymptotyczny rozkład χ^2 o $m = J - (P + Q)$ stopniach swobody.

Dla niewielkich prób **Ljung i Box** zaproponowali statystykę:

$$LB = T^2 \sum_{j=1}^J \frac{1}{T-j} \hat{\rho}_{e,j}^2$$

Test portmanteau - przykład

Wyniki testu Ljunga-Boxa dla modelu ARMA(1,2) dla inflacji CPI w Polsce:

J	LB	prawd.
4	2.393	0.12186
5	2.426	0.29736
6	3.099	0.37655
8	4.119	0.53238
10	12.437	0.08708
12	18.453	0.03027

Prognoza punktowa

Niech dany będzie proces ARMA(P, Q), dla którego znamy y_t oraz e_t dla $t \leq T$.

Prognoza punktowa jest liczona w sposób rekurencyjny:

$$E(y_{T+1}) = \alpha_0 + \alpha_1 y_T + \dots + \alpha_P y_{T+1-P} + \theta_1 e_T + \dots + \theta_Q e_{T+1-Q}$$

$$E(y_{T+2}) = \alpha_0 + \alpha_1 E(y_{T+1}) + \dots + \alpha_P y_{T+2-P} + \theta_2 e_T + \dots + \theta_Q e_{T+2-Q}$$

$$E(y_{T+3}) = \alpha_0 + \alpha_1 E(y_{T+2}) + \dots + \alpha_P y_{T+3-P} + \theta_3 e_T + \dots + \theta_Q e_{T+3-Q}$$

\vdots

$$E(y_{T+k}) = \alpha_0 + \alpha_1 E(y_{T+k-1}) + \dots + \alpha_P E(y_{T+k-P}) \text{ dla } k > P \text{ i } k > Q$$

Wariancja błędu prognozy

Postać $MA(\infty)$ modelu ARMA:

$$y_t = \mu + \epsilon_t + \psi_1 \epsilon_{t-1} + \psi_2 \epsilon_{t-2} + \dots$$

Błąd losowy prognozy:

$$y_{T+1} - E(y_{T+1}) = \epsilon_{T+1}$$

$$y_{T+2} - E(y_{T+2}) = \epsilon_{T+2} + \psi_1 \epsilon_{T+1}$$

\vdots

$$y_{T+k} - E(y_{T+k}) = \epsilon_{T+k} + \sum_{i=1}^{k-1} \psi_i \epsilon_{T+k-i}$$

Ponieważ $Cov(\epsilon_t, \epsilon_s) = 0$ oraz $Var(\epsilon_t) = \sigma_\epsilon^2$, to wariancja prognozy na k okresów wynosi:

$$\sigma_k^2 = Var(y_{T+k}) = \sigma_\epsilon^2 \left(1 + \sum_{i=1}^{k-1} \psi_i^2 \right)$$

Uwaga: nie uwzględniono innych błędów prognozy!!!

Prognoza przedziałowa

Wzór na prognozę przedziałową:

$$P(E(y_{T+k}) - c_\alpha \sigma_k \leq y_{T+k} \leq E(y_{T+k}) + c_\alpha \sigma_k) = 1 - \alpha,$$

gdzie c_α jest wartością krytyczną rozkładu normalnego dla poziomu istotności α .

