Modelowanie parametryczne

Modelowania parametryczne[1] polega na takim wyznaczeniu transmitacji filtra liniowego H(z), który upodabnia w sensie średniokwadratowym sygnał fitrowany y[n] z dodanym białym szumem do sygnału modelowanego x[n]:

$$H(z) = \frac{\sum_{k=0}^{q} b_k \cdot z^{-k}}{1 + \sum_{k=1}^{p} a_k \cdot z^{-k}} = \frac{B(z)}{A(z)} \tag{1}$$

W zależności od budowy filtru H(z) wyróżniamy modele:

- autoregresywny (AR - ang. *autoregressive*), gdy występują tylko wielomian mianownika, a licznik równa się 1;
- ze średnią ruchomą (MA - ang.*moving average*), gdy występuje tylko wielomian licznika (mianownik równa się 1);
- ARMA - gdy licznik i mianownik jednocześnie jest różny od 1.

1 Modelowanie autoregresyjne AR

Model autoregresywny (AR - ang. *autoregressive*) służy głównie do modelowania sygnałów, których widmo jest prążkowe (np. suma sygnałów harmonicznych). Model ten można zapisać w postaci:

$$x[n] = c + \sum_{k=1}^{p} \varphi_k \cdot x[n-k] + \varepsilon[n] \tag{2}$$

gdzie: c - stała, φ - parametry modelu, ε - biały szum (o średniej 0 i wariancji σ^2)

Obliczenie wartości parametrów modelu sprowadza się do rozwiązania układu równań liniowych, zapisanych w postaci macierzowej (równanie Yule'a-Walkera):

$$\begin{bmatrix} R_{xx}(0) & R_{xx}(1) & \cdots & R_{xx}(p) \\ R_{xx}(1) & R_{xx}(0) & \cdots & R_{xx}(p-1) \\ R_{xx}(2) & R_{xx}(1) & \cdots & R_{xx}(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ R_{xx}(p) & R_{xx}(p-1) & \cdots & R_{xx}(0) \end{bmatrix} \cdot \begin{bmatrix} 1 \\ \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_p \end{bmatrix} = \begin{bmatrix} \sigma^2 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{3}$$

gdzie $R_{xx}(m)$ wartość funkcji autokorelacji:

$$R_{xx}(m) = \sum_{n=0}^{N-1-|m|} x[n] \cdot x^*[n-m] \tag{4}$$

Równanie 3 można rozbić na dwie części (osobny pierwszy wiersz i reszta):

$$\begin{bmatrix} R_{xx}(0) & R_{xx}(1) & \cdots & R_{xx}(p-1) \\ R_{xx}(1) & R_{xx}(0) & \cdots & R_{xx}(p-2) \\ R_{xx}(2) & R_{xx}(1) & \cdots & R_{xx}(p-3) \\ \vdots & \vdots & \ddots & \vdots \\ R_{xx}(p-1) & R_{xx}(p-2) & \cdots & R_{xx}(0) \end{bmatrix} \cdot \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_p \end{bmatrix} = - \begin{bmatrix} R_{xx}(1) \\ R_{xx}(2) \\ R_{xx}(3) \\ \vdots \\ R_{xx}(p) \end{bmatrix} \Rightarrow \mathbf{R} \cdot \mathbf{a} = -\mathbf{r} \tag{5a}$$

$$R_{xx}(0) + \mathbf{a}^{\mathbf{T}} \cdot \mathbf{r} = \sigma^2 \tag{5b}$$

[1]Rozdział na podstawie: T. Zieliński (2009) *Cyfrowe Przetwarzanie Sygnałów*, WKŁ, Warszawa

copyright: M. Dwornik
dwornik@agh.edu.pl

2 Modelowania ze średnią ruchomą MA

Modelowanie predykcyjne z wykorzystaniem średniej ruchomej stosowane jest głównie tam, gdzie widmo ma charakter gładki. Do wyznaczenie wartości współczynników b_k z równania 1 często stosuje się modelowanie AR, przy czym ilość współczynników q licznika MA powinna być minimum 4 razy większa od ilości współczynników p mianownika MA.

Po obliczeniu wartości współczynników a, tworzona jest macierz Toeplitza $\mathbf{A}$ o rozmiarze `(q+p) x q`, gdzie wierszami zapisane są wartości współczynników a poczynając od drugiego wiersza (dla przykładu: p=4 i q=3) oraz wektor δ:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ a_1 & 1 & 0 \\ a_2 & a_1 & 1 \\ a_3 & a_2 & a_1 \\ a_4 & a_3 & a_2 \\ 0 & a_4 & a_3 \\ 0 & 0 & a_4 \end{bmatrix} \quad \delta = \begin{bmatrix} -a_1 \\ -a_2 \\ -a_3 \\ -a_4 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \tag{6}$$

$$b = (\mathbf{A}^T \cdot \mathbf{A})^{-1} \cdot \mathbf{A}^T \cdot \delta \tag{7}$$

3 ARMA

Model ten stanowi połączenie metod przedstawionych powyżej. Wartości współczynników a i b określane są w następujących krokach:

1. Wyznaczamy wartości współczynników a (dla $M > p + q$ oraz $m > q$) z równania $\mathbf{R} \cdot a = r$:

$$\begin{bmatrix} R_{xx}(q) & R_{xx}(q-1) & \cdots & R_{xx}(q-p+1) \\ R_{xx}(q+1) & R_{xx}(q) & \cdots & R_{xx}(q-p+2) \\ R_{xx}(q+2) & R_{xx}(q+1) & \cdots & R_{xx}(q-p+3) \\ \vdots & \vdots & \ddots & \vdots \\ R_{xx}(M-1) & R_{xx}(M-2) & \cdots & R_{xx}(M-p) \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} -R_{xx}(q+1) \\ -R_{xx}(q+2) \\ -R_{xx}(q+3) \\ \vdots \\ -R_{xx}(M) \end{bmatrix} \tag{8}$$

$$a = \left(\mathbf{R}^T \cdot \mathbf{R}\right)^{-1} \cdot \mathbf{R}^T \cdot r \tag{9}$$

2. Filtracja sygnału x(t) filtrem o współczynnikach a.

3. Obliczenie współczynników b modelu MA dla przefiltrowanego sygnału.

4 Econometrics toolbox

Modelowania parametryczne AR, MA i ARMA są zawarte w Econometrics toolbox. Modelowanie AR (przewidywanie) odbywa się w kilku krokach z wykorzystaniem kolejnych funkcji:

1. Stworzenie odpowiedniego modelu: `model = arima(p, D, q)`, gdzie:
   - p: ilość współczynników wielomianu autoregresji (dla AR>0, dla MA=0) ;

- D: stopień różniczkowania (zwykle 0, dla danych z liniowym trendem 1);
- q: stopień wielomianu odpowiedzialnego za średnią ruchomą (dla AR=0, dla MA>0);
wartości p, D, q muszą być nieujemnymi liczbami całkowitymi.

2. Estymacja : `est_mod = estimate(model, presample, 'Y0', estymacja)`, gdzie:
- presample - dane przedwstępne: `presample = dane(1:model.P)`;
- `'Y0'` - nazwa parametru danych przedwstępnych;
- estymacja - dane służące do estymacji: `estymacja = dane((model.P+1):N)`, gdzie N - ilość próbek wektora
3. Modelowanie: `predykcja = forecast(est_mod, N_p, dane_wej)`, gdzie
- N_p - ile danych w przód chcemy przewidzieć
- dane_wej - wektor zawierający minimum `P` ostatnich danych z wektor `estymacja`.

5 Zadania

1. Stwórz sygnały ($t = 0 - 10$s, $F_s = 100$Hz)

   - x1(t) - suma sygnałów harmonicznych (A1=1, F1=10Hz; A2=0.5, F2=3.75; A3=0.75, F3=18Hz) oraz szumu gaussowskiego (σ=0.02);
   - x2(t) - sklejenie sygnałów harmonicznych z podpunktu 1 (f1 dla t=0-3s, f2 dla t=3-6s, f3 dla t=6-9s, sygnał zerowy dla t>9s) + szum;
   - x3(t) - sygnał harmoniczny o liniowym wzroście częstotliwości f(0s)=0.5Hz, f(10s)=5Hz i stałej amplitudzie A=1;
   - x4(t) - sygnał harmoniczny o f=5Hz i A=1 i wartości NaN dla t=4s i t=7s.

   Dla każdego takiego sygnału (osobno), policz wartości współczynników dla t=<10, 10.5>s korzystając z metody AR (z i bez użycia funkcji `arima`) oraz MA i ARMA (korzystając z `arima`).

2. Pobierz dane odnośnie średniego kursu walut Narodowego Banku Polskiego (tabela A w formie XLS: https://www.nbp.pl/home.aspx?f=/kursy/arch_a.html).
Korzystając z narzędzi AR, MA i ARMA spróbuj przewidzieć kurs euro i dolara w grudniu 2023 na podstawie danych od stycznia do listopada. Dane porównaj z rzeczywistymi kursami.