

Iwona NABAGŁO

IDENTYFIKACJA STRUKTURY METODĄ GRAMA-SCHMIDTA NIELINIOWYCH MODELI CIĄGÓW CZASOWYCH

Streszczenie. Zagadnienie identyfikacji struktury modeli nieliniowych, ze względu na dużą złożoność i różnorodność modeli, napotyka na wiele problemów. Dla modeli liniowych względem parametrów często wykorzystuje się metody dobrze znane i opracowane dla modeli liniowych.

W niniejszej pracy zastosowano do identyfikacji nieliniowych modeli sygnałów losowych NARMA algorytm identyfikacji metodą najmniejszych kwadratów, oparty na ortogonalnej dekompozycji macierzy regresji metodą Grama-Schmidta i uzupełnioną o algorytm jednoczesnego doboru optymalnej struktury. Procedurę tę zastosowano do jednoczesnej estymacji parametrów i wyboru struktury modeli biliniowych.

STRUCTURE IDENTIFICATION OF NONLINEAR TIME SERIES MODELS USING GRAM-SCHMIDT METHOD

Summary. Identification algorithm based on classical Gram-Schmidt transformation which combines structure determination and parameter estimation is reviewed. This algorithm is extended for a class of discrete-time non-linear time series which polynomial models are linear in the parameters. Structure identification algorithm is then applied for bilinear time series.

1. Wstęp

W świecie rzeczywistym możemy spotkać sygnały, do opisu których modele liniowe są niewystarczające. Stosujemy wówczas modele nieliniowe. Metody stosowane do identyfikacji modeli nieliniowych zakładają znajomość struktury modelu a priori. Jednak struktura modelu jest zazwyczaj nieznaną i zachodzi konieczność doboru adekwatnej struktury modelu.

Algorytmy identyfikacji struktury zależą od rodzaju nieliniowości modelu, a także od sposobu prowadzenia doświadczenia identyfikacyjnego. Wśród modeli nieliniowych na szczególną uwagę zasługują modele liniowe względem parametrów. Przedstawienie modelu w takiej postaci umożliwia zastosowanie do identyfikacji jego parametrów metod doskonale znanych i stosowanych do identyfikacji modeli liniowych (np. rozwiązanie układu równań normalnych

Gausa, dekompozycja macierzy informacji). Jednak nadal aktualny pozostaje problem doboru struktury. W publikowanych pracach autorzy proponują modyfikację istniejących metod, tak aby pozwoliły na jednoczesną identyfikację struktury i parametrów. W artykule [5] podano ogólny algorytm do estymacji parametrów obiektu SISO metodą najmniejszych kwadratów, oparty na ortogonalnej dekompozycji macierzy regresji różnymi metodami, uzupełnionymi algorytmami doboru optymalnej struktury. Jedną z proponowanych metod jest metoda Grama-Schmidta z algorytmem doboru optymalnej struktury, którą ze względu na jakość działania i łatwość implementacji zastosowano dla modeli biliniowych.

W poniższej pracy algorytm ten zostanie przedstawiony dla nieliniowych sygnałów losowych opisanych dowolnym modelem wielomianowym, a następnie zaadaptowany do identyfikacji struktury i estymacji parametrów dla biliniowych modeli ciągów czasowych.

2. Sformułowanie problemu

Dyskretny, nieliniowy proces stochastyczny możemy opisać ogólnym modelem NARMA (Non-linear AutoRegressive Moving Average).

$$y(t) = f [y(t-1), \dots, y(t-n_y), e(t-1), \dots, e(t-n_e)] + e(t), \quad (1)$$

gdzie:

n_y, n_e - maksymalne przesunięcia zmierzonego sygnału i szumu;

$e(t)$ - „biały szum” o zerowej wartości oczekiwanej;

$f(\cdot)$ - dowolna nieliniowa funkcja.

Podklasą modeli NARMA są modele typu NAR (Non-linear AutoRegressive):

$$y(t) = f [y(t-1), \dots, y(t-n_y)] + e(t). \quad (2)$$

Większość dalszej części rozważań będzie dotyczyła tej klasy modeli. Rozszerzenie na modele NARMA zostanie przedstawione w osobnym rozdziale.

Zazwyczaj postać nieliniowej funkcji $f(\cdot)$ jest nieznaną. Jednakże każdą nieliniową funkcję ciągłą można w miarę dokładnie aproksymować modelem wielomianowym. Przyjmując funkcję $f(\cdot)$ jako wielomian stopnia l otrzymujemy następującą postać modelu:

$$y(t) = \theta_0 + \sum_{i_1=1}^{n_y} \theta_{i_1} x_{i_1}(t) + \sum_{i_1=1}^{n_y} \sum_{i_2=i_1}^{n_y} \theta_{i_1 i_2} x_{i_1}(t) x_{i_2}(t) + \dots$$

$$\dots + \sum_{i_1=1}^{n_y} \dots \sum_{i_l=i_{l-1}}^{n_y} \theta_{i_1 \dots i_l} x_{i_1}(t) \dots x_{i_l}(t) + \epsilon(t), \quad (3)$$

gdzie: $x_1(t) = y(t-1)$, $x_2(t) = y(t-2)$, ..., $x_{n_y}(t) = y(t-n_y)$.

Model (3) można przedstawić jako liniowy model regresyjny o postaci:

$$y(t) = \sum_{i=1}^M \varnothing_i(t) \theta_i + \epsilon(t), \quad (4)$$

gdzie:

$$t = 1, \dots, N \quad M = \frac{(n_y + 1)!}{n_y! 1!}$$

N - liczba pomiarów $y(t)$;

$\varnothing_i(t)$ - iloczyny $x_1(t)$ do $x_{n_y}(t)$ aż do stopnia l ;

$\varnothing_i(t) = 1$ - odpowiada wartości stałej;

$\epsilon(t)$ - błąd modelowania + błąd stochastyczny $\epsilon(t)$;

θ_i - nieznanne parametry.

Postać wektorowa modelu jest następująca:

$$y = \Phi \Theta + \Xi, \quad (5)$$

gdzie:

$$y = \begin{bmatrix} y(1) \\ \vdots \\ y(N) \end{bmatrix}, \quad \Theta = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_M \end{bmatrix}, \quad \Xi = \begin{bmatrix} \epsilon(1) \\ \vdots \\ \epsilon(N) \end{bmatrix}$$

$$\Phi = \begin{bmatrix} \phi_1(1) & \dots & \phi_M(1) \\ \vdots & \ddots & \vdots \\ \phi_1(N) & \dots & \phi_M(N) \end{bmatrix} = [\phi_1 \phi_2 \dots \phi_M]$$

Wymiar macierzy: $\dim \Phi = N \times M$; $M \leq N$.

Możliwość przedstawienia równania modelu nieliniowego w postaci modelu liniowego względem parametrów sugeruje możliwość wykorzystania do identyfikacji metod opracowanych dla modeli liniowych, np. metody najmniejszych kwadratów (NK).

Jeżeli struktura wielomianu jest znana a priori, tzn. macierz regresji Φ zawiera tylko istotne składniki wielomianu, wówczas wystarczy znaleźć metodą NK oceny parametrów $\hat{\Theta}$.

Jeżeli macierz Φ będzie reprezentowała pełny model, to problem jednoczesnej identyfikacji struktury i estymacji parametrów można sformułować następująco:

Wyznacz macierz Φ , (zawierającą tylko istotne składniki) z macierzy Φ i znajdź odpowiadającą jej oceny $\hat{\Theta}$.

Ponieważ liczba wszystkich możliwych składników modelu jest zazwyczaj bardzo duża, trudno znaleźć optymalne rozwiązanie powyższego problemu. Przykładowo dla $n_y = 2$, $l = 2$ liczba nieznanych parametrów wynosi:

$$M = \frac{(n_y + 1)!}{n_y! 1!} = \frac{(2 + 2)!}{2! 2!} = 6.$$

Dla $n_y = 4$, $l = 2$ liczba parametrów $M = 15$, ale dla $n_y = 10$, $l = 5$ liczba parametrów wynosi już $M = 3003$.

3. Klasyczna metoda Grama-Schmidta

Rozwiązanie problemu NK sprowadza się do rozwiązania układu równań normalnych Gaussa:

$$(\Phi^T \Phi) \Theta = \Phi^T y, \quad (6)$$

gdzie: $\Phi^T \Phi$ jest nazwana macierzą informacji. Najlepsze parametry $\hat{\Theta}$ możemy wyznaczyć stosując ortogonalną dekompozycję macierzy Φ .

Aby zabezpieczyć się przed słabym uwarunkowaniem macierzy informacji, dokonujemy faktoryzacji w następujący sposób:

$$\Phi = WA, \quad (7)$$

gdzie

$$A = \begin{bmatrix} 1 & \alpha_{12} & \alpha_{13} & \cdots & \alpha_{1M} \\ & 1 & \alpha_{23} & \cdots & \alpha_{2M} \\ & & \ddots & \ddots & \vdots \\ & & & 1 & \alpha_{M-1M} \\ & & & & 1 \end{bmatrix} \quad (8)$$

jest górno-trójkątną macierzą o wymiarach: $\dim A = M \times M$,

$$W = \begin{bmatrix} w_1(1) & \dots & w_M(1) \\ \vdots & \ddots & \vdots \\ w_1(N) & \dots & w_M(N) \end{bmatrix} = [w_1 \dots w_M] \quad (9)$$

jest macierzą o wymiarach: $N \times M$ i ortogonalnych kolumnach.

$$W^T W = D, \quad (10)$$

gdzie D jest macierzą diagonalną, dodatnio określoną.

W metodzie Grama-Schmidta (GS) ortogonalizacja macierzy Φ przebiega następująco:

W k -tym kroku ortogonalizowana jest k -ta kolumna względem poprzednio zortogonalizowanych $k - 1$ kolumn; operację powtarza się dla $k = 2, \dots, M$:

$$\left. \begin{aligned} w_1 &= \varnothing_1 \\ \alpha_{ik} &= \frac{w_i^T \varnothing_k}{w_i^T w_i}, \quad 1 \leq i < k \\ w_k &= \varnothing_k - \sum_{i=1}^{k-1} \alpha_{ik} w_i \end{aligned} \right\} k = 2, \dots, M \quad (11)$$

$$w_i^T \varnothing_k = \sum_{t=1}^N w_i(t) \varnothing_k(t) \quad (12)$$

Definiując i przekształcając:

$$\left. \begin{aligned} \Phi^T \Phi \Theta &= \Phi^T y \\ \Phi &= W A \end{aligned} \right\} \Rightarrow \left. \begin{aligned} A \Theta &= (W^T W)^{-1} W^T y \\ W^T W &= D \end{aligned} \right\} \Rightarrow A \Theta = D^{-1} W^T y$$

oraz odpowiednio podstawiając:

$$g = D^{-1} W^T y \Rightarrow g_i = \frac{w_i^T y}{w_i^T w_i}, \quad i = 1, \dots, M \quad (13)$$

otrzymujemy zależność:

$$A \Theta = g. \quad (14)$$

Oceny parametrów $\hat{\Theta}$ wyznaczamy z poniższej zależności, stosując algorytm obliczeń wstecz:

$$\hat{\Theta} = A^{-1} g. \quad (15)$$

4. Wybór składników struktury

Naszym zadaniem jest wyselekcjonowanie z pełnej macierzy Φ macierzy Φ_s , zawierającej kolumny odpowiadające tylko istotnym składnikom modelu. Macierz Φ_s ma M_s ($M_s < M$, $M_s \leq N$) kolumn.

Postać wektorowa modelu zawierającego tylko istotne składniki będzie następująca:

$$y = \Phi_s \Theta_s + \Xi \quad (16)$$

Faktoryzacja macierzy Φ_s przebiega tak samo jak poprzednio:

$$\Phi_s = W_s A_s, \quad (17)$$

gdzie W_s jest macierzą o wymiarach $N \times M_s$, która zawiera M_s ortogonalnych kolumn, a macierz A_s o wymiarach $M_s \times M_s$ jest macierzą górno-trójkątną.

Dokonując przekształceń jak dla pełnej macierzy Φ otrzymamy wzór (14) w postaci:

$$A_s \Theta_s = g_s. \quad (18)$$

Z równań (17) i (18) wynika:

$$\Phi_s \Theta_s = W_s g_s. \quad (19)$$

Równanie (16) przyjmuje postać:

$$y = W_s g_s + \Xi. \quad (20)$$

Błędy identyfikacji (residua) będą więc określone zależnością:

$$\hat{\Xi} = y - W_s g_s. \quad (21)$$

Korzystając z powyższej zależności, wariancję sygnału Y możemy opisać równaniem:

$$\frac{1}{N} (y^T y) = \frac{1}{N} \sum_{i=1}^{M_s} g_i^2 (w_{S_i}^T w_{S_i}) + \frac{1}{N} (\hat{\Xi}^T \hat{\Xi}). \quad (22)$$

Ze wzoru (22) wynika, że maksymalną wariancję błędu identyfikacji otrzymamy wówczas, gdy $M_s = 0$ (model nie zawiera żadnego składnika). Włączając i -ty składnik do modelu zmniejszamy wariancję błędu identyfikacji.

Dla każdej części struktury modelu możemy wyznaczyć stopień przynależności składnika w_i do struktury, zdefiniowany następująco:

$$\omega_i = \frac{g_i^2 (w_i^T w_i)}{y^T y}. \quad (23)$$

Im większy będzie stopień przynależności danego składnika, tym mniejszą wariancję błędu identyfikacji otrzymamy po dołączeniu tegoż składnika do struktury. Sposób selekcji macierzy Φ_s z macierzy Φ powinien więc być następujący:

Obliczając w i -tym kroku stopień przynależności dla $M - i + 1$ kolumn macierzy W wybieramy tę kolumnę w_j , dla której stopień przynależności jest największy. Kolumna ta staje się i -tą kolumną macierzy W_s .

4.1. Zmodyfikowany algorytm CGS

Algorytm ortogonalizacji macierzy Φ , połączony z wyborem struktury modelu, będzie następujący:

Pierwszy krok: dla $i = 1, \dots, M$ przyjmujemy:

$$w_i^{(1)} = \emptyset_i$$

i obliczamy dla każdej kolumny stopień przynależności $\omega_i^{(1)}$:

$$g_i^{(1)} = \frac{w_i^{(1)T} y}{w_i^{(1)T} w_i^{(1)}},$$

$$\omega_i^{(1)} = \frac{\left(g_i^{(1)}\right)^2 w_i^{(1)T} w_i^{(1)}}{y^T y}.$$

Wybieramy największy stopień przynależności:

$$\omega_j^{(1)} = \max\left\{\omega_i^{(1)}, 1 \leq i \leq M\right\}.$$

Odpowiadającą mu kolumna

$$w_{s_1} = w_j^{(1)} = \emptyset_j$$

zostaje wybrana jako pierwsza kolumna macierzy W_s razem z pierwszym elementem wektora

$$g_s, g_{s_1} = g_j^{(1)}, i \quad \omega_{s_1} = \omega_j^{(1)}.$$

Drugi krok: dla $i = 1, \dots, M$ i $i \neq j$ obliczamy:

$$\alpha_{12}^i = \frac{w_{s_1}^T \varnothing_i}{w_{s_1}^T w_{s_1}},$$

$$w_i^{(2)} = \varnothing_i - \alpha_{12}^i w_{s_1},$$

$$g_i^{(2)} = \frac{w_i^{(2)T} y}{w_i^{(2)T} w_i^{(2)}},$$

$$\omega_i^{(2)} = \frac{\left(g_i^{(2)}\right)^2 w_i^{(2)T} w_i^{(2)}}{y^T y}.$$

Wybieramy:

$$\omega_k^{(2)} = \max\left\{\omega_i^{(2)}, 1 \leq i \leq M, i \neq j\right\}.$$

Następnie

$$w_{s_2} = w_k^{(2)} = \varnothing_k - \alpha_{12} w_{s_1}$$

zostaje wybrana jako druga kolumna macierzy W_s razem z drugą kolumną macierzy A_s , $\alpha_{12} = \alpha_{12}^k$ oraz drugim elementem wektora g_s , $g_{s2} = g_k^{(2)}$, i $\omega_{s2} = \omega_k^{(2)}$.

Następne kroki są analogiczne do kroku drugiego.

Wybór modelu będzie prowadzony dopóty, dopóki nie zostanie spełnione kryterium zakończenia obliczeń, np. zgodnie z kryterium informacyjnym Akaikego AIC:

$$AIC[1] = N \ln C(\hat{\Theta}_s) + 2M_s,$$

gdzie

$$C(\hat{\Theta}_s) = \frac{1}{N} \sum_{t=1}^N \varepsilon^2(t).$$

Jeżeli:

$$AIC[1] > AIC[1-1],$$

procedura selekcji zostaje zakończona i $(1-1)$ -sza struktura zostaje wybrana.

Oceny parametrów $\hat{\Theta}_s$ wyznacza się w prosty sposób z zależności:

$$A_s \hat{\Theta}_s = g_s$$

zgodnie z algorytmem obliczeń wstecz:

$$\hat{\theta}_M = \frac{g_M}{\alpha_{MM}}$$

$$\hat{\theta}_i = \frac{g_i - \sum_{k=i+1}^M \alpha_{ik} \hat{\theta}_k}{\alpha_{ii}}, i = M-1, \dots, 1$$

Ponieważ $\alpha_{ii} = 1$, więc:

$$\hat{\theta}_M = g_M$$

$$\hat{\theta}_i = g_i - \sum_{k=i+1}^M \alpha_{ik} \hat{\theta}_k, i = M-1, \dots, 1,$$

gdzie:

$$M = M_S, \hat{\theta} = \hat{\theta}_S, g = g_S.$$

5. Rozszerzenie algorytmu na modele NARMA

Niech macierz Φ reprezentuje pełny model, podzielony na dwie macierze:

$$\Phi = [\Phi_p : \Phi_n],$$

gdzie:

Φ_p - zawiera tylko zmierzone wartości sygnału i ich iloczyny,

Φ_n - zawiera błędy modelowania i iloczyny błędów modelowania i wartości sygnału.

Zadanie jednoczesnej estymacji parametrów i identyfikacji struktury rozwiązuje się tak jak poprzednio. Należy wydzielić macierz Φ_s z macierzy Φ i znaleźć odpowiadające jej parametry. Tym razem jednak macierz Φ_s składa się z dwóch macierzy:

$$\Phi_s = [\Phi_{ps} : \Phi_{ns}], \quad (24)$$

przy czym macierz Φ_n jest nieznaną (zawiera nieznanne błędy modelowania).

Sposób postępowania powinien więc być następujący:

1. Krok pierwszy

- Należy wybrać M_{ps} kolumn macierzy Φ_{ps} z macierzy Φ zgodnie z algorytmem wyboru kolumn podanym poprzednio.
- Metodą obliczeń wstecz wyznaczyć początkowy wektor parametrów $\hat{\theta}_s^0$.

- (c) Na podstawie otrzymanego wektora parametrów $\hat{\Theta}_s^0$ należy wyznaczyć początkowy wektor błędów modelowania $\{\epsilon^0(t)\}$, utworzyć macierz Φ_n i wybrać z niej M_n kolumn macierzy Φ_n .
- (d) Do macierzy Φ_p zostają dołożone kolumny macierzy Φ_n . Korzystając z tak utworzonej pełnej macierzy należy wyznaczyć nowy wektor parametrów $\hat{\Theta}_s^1$ i wygenerować nowy wektor błędów modelowania $\{\epsilon^1(t)\}$.

2. Krok następny (ogólnie)

W k -tej iteracji dzięki wektorowi błędów modelowania z poprzedniej iteracji $\{\epsilon^{k-1}(t)\}$ tworzona jest macierz Φ_n^k . Z macierzy:

$$[\Phi_p : \Phi_n^k]$$

wybierana jest macierz Φ_s^k :

$$\Phi_s^k = \begin{bmatrix} \Phi_{p_s}^k : \Phi_{n_s}^k \end{bmatrix}.$$

Wyznaczany jest wektor parametrów $\hat{\Theta}_s^k$ i generowany nowy wektor błędów modelowania $\{\epsilon^k(t)\}$.

3. Procedura z punktu 2 jest powtarzana dopóki nie zostanie spełnione kryterium AIC.

Ostatecznie zostaje wybranych:

$$M_s = M_{p_s}^k + M_{n_s}^k$$

kolumn macierzy Φ_s^k i wygenerowany wektor ocen parametrów $\hat{\Theta}_s^k$.

5.1. Modele biliniowe

Modele biliniowe (BARMA) należą do klasy modeli nieliniowych względem zmiennych, ale liniowych względem parametrów. Ogólna postać modelu biliniowego jest następująca:

$$A(D)y(t) = C(D)\epsilon(t) + \sum_{k=1}^K \sum_{l=1}^L \beta_{kl} \epsilon(t-k)y(t-l),$$

gdzie:

$$A(z^{-1}) = 1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_{d_A} z^{-d_A},$$

$$C(z^{-1}) = 1 + c_1 z^{-1} + c_2 z^{-2} + \dots + c_{d_C} z^{-d_C}.$$

Macierze Φ_p i Φ_n dla modelu biliniowego mają następującą postać:

$$\Phi_p = [1 \quad y(t-1) \quad y(t-2) \quad \dots \quad y(t-dA)]$$

$$\Phi_n = [\begin{array}{cccc} e(t-1) & e(t-2) & \dots & e(t-dC) \\ e(t-1)y(t-1) & e(t-1)y(t-2) & \dots & e(t-1)y(t-L) \\ e(t-2)y(t-1) & e(t-2)y(t-2) & \dots & e(t-2)y(t-L) \\ \vdots & \vdots & \ddots & \vdots \\ e(t-K)y(t-1) & e(t-K)y(t-2) & \dots & e(t-K)y(t-L) \end{array}]$$

Algorytm identyfikacji struktury modeli BARMA

1. W pierwszym kroku szukamy dla ciągu biliniowego najlepszego modelu liniowego typu AR, stosując kryterium AIC. Z macierzy Φ_p wyznaczamy macierz Φ_{pi} i początkowy wektor parametrów $\hat{\Theta}_S^0$. Używając otrzymanego wektora parametrów generujemy początkowy wektor błędów modelowania $\{\varepsilon^0(t)\}$ i tworzymy macierz Φ_n .
2. Dalszy sposób postępowania jest analogiczny do algorytmu podanego dla modeli NARMA.

5.2. Badania symulacyjne

Zasymulowano ciągi biliniowe o postaci:

1. $y(t) = 0.45 y(t-3) e(t-2) + e(t)$
2. $y(t) = 0.45 y(t-2) e(t-3) + e(t)$
3. $y(t) = -0.3 y(t-1) + 0.45 y(t-1) e(t-1) + e(t)$,

gdzie $e(t)$ jest „białym szumem” o wariancji $\lambda^2 = 1$.

Przeprowadzono następujące doświadczenie identyfikacyjne: zasymulowano 100 ciągów o liczbie pomiarów 1024 każdy, a następnie przeprowadzono zgodnie z podanym algorytmem jednoczesną identyfikację struktury i parametrów dla każdego ciągu.

Uzyskano następujące wyniki (uśrednione):

Numer ciągu	Estymaty parametrów
1	$\bar{\beta}_{32} = 0.4436 \quad \bar{\beta}_{23} = 0,0071 \quad \bar{\beta}_{kl} = r10^{-5}$
2	$\bar{\beta}_{23} = 0.4436 \quad \bar{\beta}_{32} = 0,0041 \quad \bar{\beta}_{kl} = r10^{-5}$
3	$\bar{a}_1 = -0.306 \quad \bar{\beta}_{11} = 0,3832 \quad \bar{\beta}_{kl} = r10^{-5}$

Podczas identyfikacji struktury dla zaledwie kilku realizacji ciągów została znaleziona struktura modelu odpowiadająca dokładnie strukturze obiektu. Dla większości przykładów algorytm wybierał dodatkowe składniki struktury (różne), ale wyznaczone dla nich wartości parametrów były rzędu co najwyżej 10^{-2} . Parametry te zostały uwzględnione w tabeli w postaci ogólnej $\bar{\beta}_{kl} = r10^{-5}$.

LITERATURA

1. Bielińska E., Minimum variance prediction of bilinear time series - direct and adaptive version. *Journal of Forecasting*, Vol.12, pp. 459-280, 1993
2. Billings S., Leontaritis I. J., Parameter estimation techniques for nonlinear systems. 6-th IFAC Symposium on Identification and System Parameters Estimation, Arlington, Virginia, USA, 1982
3. Box G.E., Jenkins G.M., Time series analysis. Holden Day, San Francisco, 1970
4. Chen S., Billings S.A., Modelling and analysis of nonlinear time series. *Int. J. Control*, Vol. 50, Nr 6, pp. 2151-2171, 1989
5. Chen S., Billings S.A., Orthogonal least squares methods and their application to non-linear system identification. *Int. J. Control*, Vol. 50, Nr 5, pp. 1873-1896, 1989
6. Haber R., Unbehauen H., Structure Identification of Nonlinear Dynamic Systems - A Survey on Input / Output Approaches. *Automatica*, Vol. 26, Nr 4, pp. 651-677, 1990
7. Korenberg M., Billings S.A., Liu Y.P., McIlroy P.J., Orthogonal parameter estimation algorithm for non-linear stochastic systems. *Int. J. Control*, Vol. 48, Nr 1, pp. 193-210, 1988
8. Ljung L., System Identification: Theory for the User. Prentice-Hall, New Jersey, 1987
9. Tang Z., Mohler R.R., Bilinear time series: Theory and application. *Lecture Notes in Control and Information Sciences*, Vol. 106, Springer - Verlag, Berlin, 1989

Recenzent: Doc. dr hab. Zbigniew Nahorski

Wpłynęło do Redakcji 23.06.1994 r.

Abstract

Most signals encountered in the real world are non-linear. We can describe them using a polynomial time series model linear in the parameters, e.g. NARMA model (Non-linear AutoRegressive Moving Average).

When the model's structure is known, only the values of the parameters should be identified. Identification can thus be formulated as a standard least squares problem which can be solved using various well-developed numerical techniques. Unfortunately the model structure of real systems is rarely known a priori.

The model structure is generally unknown at the beginning of the identification and we often have to consider the full model set. In reality each model may involve only a few significant terms which adequately characterize the system dynamics. The determination of the structure or which terms to include in the final model is essential.

Let Φ represents the full model matrix. The combined problem of structure selection and parameter estimation can be define as follows:

Select a submatrix Φ_s (which contains only significant terms) of matrix Φ and find the corresponding parameter estimate $\hat{\Theta}_s$.

Because the number of all the possible terms can easily became excessively large it is very difficult to attain the optimal solution of the above problem.

There are orthogonal algorithms which efficiently combine structure selection and parameters estimation. These methods base on the well-known techniques of orthogonal decomposition of the regression matrix.

Identification algorithm based on orthogonal least squares method, know as classical Gram-Schmidt transformation is reviewed. This algorithm is extended for a class of discrete-time non-linear time series which polynomial models are linear in the parameters. Structure identification algorithm is then applied for bilinear time series.