VIII KRAJOWE
SYMPOZJUM
NAUK RADIOWYCH
WROCŁAW,
15 - 16 lutego 1996 r.

Włodzimierz Greblicki
Instytut Cybernetyki Technicznej
Politechnika Wrocławska

Nieparametryczna Identyfikacja Systemów Dynamicznych

STRESZCZENIE

W pracy przedstawiono zadania nieparametrycznej identyfikacji systemów w sytuacji, gdy system pobudzany jest sygnałem losowym, a obserwacje zakłócone są przypadkowym szumem. Omówiono zagadnienie identifikacji nieliniowej charakterystyki zwykłego systemu statycznego, a następnie przedstawiono zadania nieparametrycznej identyfikacji nieliniowych, dynamicznych systemów Hammersteina i Wienera. Podano algorytmy identyfikacji podsystemów nieliniowych. Omówiono problem ich zbieżności

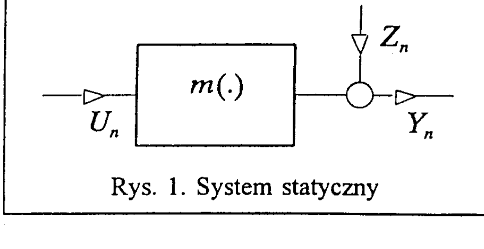
WSTĘP

W procesie identyfikacji systemów korzysta się z dwóch rodzajów informacji. Pierwsza jest natury empirycznej, druga jest informacją wstępną o systemie. Dysponuje się nią już przed przystąpieniem do wykonywania jakichkolwiek pomiarów. Informację empiryczną tworzą wyniki pomiarów sygnałów w systemie. Najczęściej są one wykonane na jego wejściu i wyjściu. Odnośnie informacji wstępnej, zakłada się zazwyczaj, że ma ona charakter parametryczny, co oznacza, że opis systemu znany jest z dokładnością do skończonej liczby parametrów. Celem identyfikacji jest wtedy wyznaczenie lub estymacja nieznanych wartości tych parametrów na podstawie danych pomiarowych. W poniższej pracy informacja wstępna o badanym systemie jest uboższa od wspomnianej powyżej, początkowy brak informacji o systemie jest dalej posunięty. W rezultacie zadanie identyfikacji systemu jest nieparametryczne. Poniżej przedstawione zostaną algorytmy identyfikacji dwóch rodzajów nieliniowych systemów dynamicznych o złożonej strukturze najczęściej badanych w interesującym nas kontekście, a mianowicie systemu Hammersteina i Wienera. Najpierw jednak, w celu zaprezentowania podstawowych idei, pokażemy jak estymować nieparametrycznie charakterystykę elementu statycznego. Podamy także najważniejsze własności tych algorytmów.

Warto dodać, że identyfikacja systemów o złożonej strukturze jest przedmiotem zainteresowania wielu autorów, patrz np. [1], [2]. Spowodowane to jest dużym zapotrzebowaniem w różnych obszarach działania praktycznego. Tutaj warto wspomnieć o zastosowaniach w telekomunikacji, [3].

SYSTEM STATYCZNY

W celu przedstawienia najważniejszych problemów związanych z nieparametryczną identyfikacją systemów zajmiemy się na początek systemem statycznym przedstawionym na Rys. 1. W systemie tym $Y_n = m(U_n) + Z_n$, gdzie sygnał wejściowy $\{U_n; n = ...,-1,0,1,... \}$ jest stacjonarnym losowym białym szumem, Z_n przypadkowym zakłóceniem o zerowej wartości oczekiwanej, natomiast m nieznaną charakterystyką, którą należy identyfikować na podstawie par pomiarów (U_0,Y_0), (U_1,Y_1), ...,(U_n,Y_n). O charakterystyce tej zakładamy jedynie, że jest ona funkcją mierzalną w sensie Borela.



Rys. 1. System statyczny

Jest to niezbędne dla opisania zadania w języku probabilistyki. Przedstawiając podany poniżej algorytm identyfikacji wychodzi się ze spostrzeżenia, że m można przedstawić jako funkcję regresji, gdyż $m(u) = E\{Y_n | U_n = u\}$.

Zauważmy, że $m(u) = g(u)/f(u)$, gdzie f jest gęstością prawdopodobieństwa sygnału wejściowego, natomiast $g(u) = f(u)E\{Y_n | U_n = u\}$. Jako przybliżenia g oraz f można przyjąć odpowiednio

$$\frac{1}{2h}\int_{u-h}^{u+h}\left[\int y f(v,y)\,dy\right]dv \quad \text{oraz} \quad \frac{1}{2h}\int_{u-h}^{u+h} f(v)\,dv,$$

gdzie h jest pewną dodatnią liczbą. Przez $f(.,.)$ oznaczyliśmy tutaj łączną gęstość prawdopodobieństwa wejścia i wyjścia. Jest oczywiste, że gdy h maleje do zera, powyższe wyrażenia zbiegają się odpowiednio do $g(u)$ i $f(u)$. Dają się one łatwo oszacować na podstawie pomiarów, bowiem jako nieobciążone estymatory podanych wielkości można przyjąć

$$\frac{1}{2nh}\sum_{i=1}^{n} Y_i K\left(\frac{u-U_i}{h}\right) \quad \text{oraz} \quad \frac{1}{2nh}\sum_{i=1}^{n} K\left(\frac{u-U_i}{h}\right),$$

gdzie $K(u)$ jest równe 1/2 dla u należącego do przedziału (-1,1) oraz równe zero poza tym przedziałem. Estymator nieznanej charakterystyki m można zatem zdefiniować następująco:

$$\hat{m}(u) = \sum_{i=1}^{n} Y_i K\left(\frac{u-U_i}{h(n)}\right) \Big/ \sum_{i=1}^{n} K\left(\frac{u-U_i}{h(n)}\right),$$

gdzie $\{h(n)\}$ jest dodatnim ciągiem liczbowym. Jako jądro estymatora przyjmuje się przy tym także funkcje inne niż podana powyżej. Można wykazać, że prawdziwe jest podane poniżej twierdzenie.

Twierdzenie 1. Załóżmy, że $\int m^2(u)f(u)\,du < \infty$ oraz $EZ_0^2 < \infty$. Niech K będzie nieujemną, ograniczoną funkcją taką, że $\int K(u)\,du < \infty$ oraz $\lim_{|u|\to\infty} |u| K(u) = 0$. Jeżeli

$$h(n) \xrightarrow{n} 0 \quad \text{oraz} \quad nh(n) \xrightarrow{n} \infty,$$

to

$$\hat{m}(u) \xrightarrow{n} m(u) \quad \text{według prawdopodobieństwa}$$

w każdym punkcie u, w którym zarówno charakterystyka m, jak i gęstość f są ciągłe oraz $f(u) > 0$.

Wynik powyższy, a także podobne, można znaleźć w wielu pracach, patrz np. [7] lub [6]. Pierwsza z tych pozycji jest wyczerpującą monografią, druga ma charakter przeglądowy. Należy tutaj zwrócić uwagę, że istnieje wiele ciągów spełniających warunki twierdzenia. Wybrać możemy bowiem np. $h(n) = cn^{-\alpha}$, gdzie $c > 0$ oraz $0 < \alpha < 1$. Bogaty jest także wybór funkcji K. Oprócz już podanej można zastosować np. następujące: $e^{-|u|}$, e^{-u^2}, $1/(1+u^2)$.

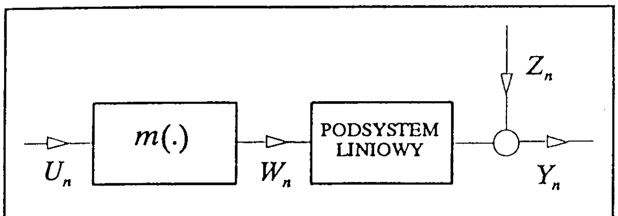
Trzeba jednak zwrócić przede wszystkim uwagę na to, że algorytm identyfikacji zbiega się dla bardzo szerokiej klasy charakterystyk. Jedyne założenie jej dotyczące nie stanowi bowiem w istocie żadnego ograniczenia z punktu widzenia zastosowań. Można więc stwierdzić, że przedstawiony algorytm można skutecznie stosować w sytuacji, gdy informacji wstępna, czyli informacja *a'priori*, jest niewielka, niemal żadna.

Zagadnienia szybkości zbieżności nie będziemy tutaj szerzej omawiać. Ograniczymy się do stwierdzenia, że dla dwukrotnie różniczkowalnych m i f interesująca nas szybkość zbieżności posiada rząd $O(n^{-2/5})$.

SYSTEM HAMMERSTEINA

Znacznie ważniejszy z punktu widzenia zastosowań, a jednocześnie trudniejszy dla analizy teoretycznej, jest problem identyfikacji nieliniowej części w dynamiczym systemie Hammersteina, Rys. 2. System ten składa się ze statycznego podsystemu nieliniowego oraz liniowego, dynamicznego. Sygnał wejściowy jest stacjonarnym białym szumem. Zakłócenie Z_n ma zerową wartość oczekiwaną. W systemie $W_n = m(U_n)$, gdzie m jest nieznaną charakterystyką. Spełnia ona następującą nierówność: $|m(u)| = c_1 + c_2|u|$, gdzie c_1 i c_2 są pewnymi, nieznanymi liczbami.

O dynamicznym systemie liniowym wiadomo jedynie, że jest asymptotycznie stabilny. Zadanie identyfikacji podsystemu nieliniowego polega na estymacji jego charakterystyki m na podstawie pomiarów wejścia i wyjścia (U_0,Y_0), (U_1,Y_1), ...,(U_n,Y_n) całego systemu. Faktem podstawowym dla konstrukcji algorytmu identyfikacji jest to, że $E\{Y_{n+1}|U_n\} = \alpha m(u)$, gdzie α jest pewną nieznaną liczbą. Ze względu na to, że sygnał łączący podsystemy nie jest dostępny pomiarowi, stałej tej nie można wyliczyć ani estymować. Problem identyfikacji polega zatem, podobnie jak w poprzednim paragrafie, na estymacji funkcji regresji. Jako algorytm identyfikacji można zatem przyjąć:



Rys. 2. System Hammersteina

$$\bar{m}(u) = \sum_{i=1}^{n} Y_{i+1} K\left(\frac{u-U_i}{h(n)}\right) \Big/ \sum_{i=1}^{n} K\left(\frac{u-U_i}{h(n)}\right),$$

Trzeba tutaj zwrócić uwagę na istotną różnicę pomiędzy sytuacją omówioną poprzednio, a obecną. Wtedy wnioskowaliśmy na podstawie obserwacji stochastycznie niezależnych par (U_0,Y_0), (U_1,Y_1), ...,(U_n,Y_n). Teraz pary $(U_0,Y_1),(U_1,Y_2),...,(U_n,Y_{n+1})$, na podstawie których wnioskujemy, są zależne, gdyż zmienne losowe $Y_1,Y_2,...,Y_{n+1}$ są zależne, ponieważ są wyjściem systemu dynamicznego. Warunki w jakich stosujemy teraz algorytm są więc trudniejsze. Można wykazać, że prawdziwy jest następujący wynik, patrz [6]:

Twierdzenie 2. Załóżmy, że $\int m^2(u)f(u)du < \infty$ oraz $EZ_0^2 < \infty$. Niech K będzie nieujemną, ograniczoną funkcją taką, że $\int K(u)du < \infty$ oraz $\lim_{|u|\to\infty} |u|K(u) = 0$. Jeżeli

$$h(n) \xrightarrow{n} 0 \quad \text{oraz} \quad nh(n) \xrightarrow{n} \infty,$$

to

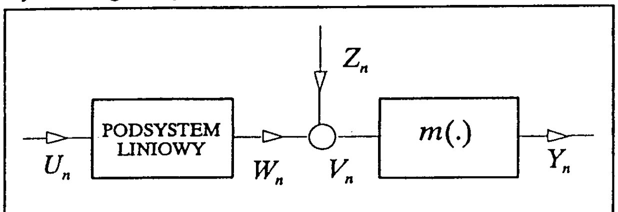
$$\bar{m}(u) \xrightarrow{n} \alpha m(u) \quad \text{według prawdopodobieństwa}$$

w każdym punkcie u, w którym zarówno charakterystyka m, jak i gęstość f są ciągłe oraz $f(u) > 0$.

Rząd szybkości zbieżności algorytmu jest taki sam jak przy identyfikacji systemu statycznego.

SYSTEM WIENERA

Kolejnym systemem, któremu poświęcimy uwagę jest system Wienera przedstawiony na Rys. 3. Jego część liniowa o nieznanym opisie jest asymptotycznie stabilna. O nieznanej charakterystyce m wiadomo jedynie, że jest różniczkowalna i jest funkcją odwracalną, tzn. że istnieje m^{-1}. Sygnał wejściowy jest gaussowskim, białym szumem, a zakłócenie o skończonej wariancji ma zerową wartość oczekiwaną. Można sprawdzić, że $E\{U_{n-1}|Y_n=y\} = \lambda m^{-1}(y)$, gdzie λ jest nieznaną i niemożliwą ani do wyliczenia ani estymacji stałą. Estymacja nieznanej charakterystyki, a raczej jej odwrotności, jest teraz równoważna estymacji funkcji regresji. Jako algorytm identyfikacji przyjmiemy teraz



Rys. 3. System Wienera

$$\bar{\mu}(u) = \sum_{i=1}^{n} U_i K\left(\frac{y-Y_{i+1}}{h(n)}\right) \Big/ \sum_{i=1}^{n} K\left(\frac{y-Y_{i+1}}{h(n)}\right),$$

Jak już wspomnieliśmy estymuje on $\lambda m^{-1}(y)$. Warunki w jakich teraz przebiega proces identyfikacji są jeszcze trudniejsze gdyż zależne obserwacje $Y_1,Y_2,...,Y_{n+1}$ są teraz związane z argumentem funkcji regresji. Warto tutaj zauważyć, że wyliczając pseudoinwersję estymatora odwrotności m^{-1}

można skutecznie estymować $m(v/\lambda)$, λ nadal nieznane. Można wykazać, że prawdziwe jest twierdzenie podane poniżej.

Twierdzenie 3. Załóżmy, że m jest jest funkcją o ograniczonej pochodnej i jest odwracalna. Niech K będzie nieujemną, ograniczoną funkcją taką, że $\int K(u)du < \infty$ oraz $\lim_{|u| \to \infty} |u| K(u) = 0$. Jeżeli

$$h(n) \xrightarrow{n} 0 \quad \text{oraz} \quad nh^2(n) \xrightarrow{n} \infty,$$

to

$$\hat{\mu}(y) \xrightarrow{n} \lambda m^{-1}(y) \quad \text{według prawdopodobieństwa}$$

w prawie każdym punkcie y.

Zauważmy, że jeśli jako $h(n) = cn^{-\alpha}$, to założenia twierdzenia są spełnione dla $0 < \alpha < 1/2$. Dla charakterystyki podobnej jak w poprzednich dwóch przypadkach, algorytm jest teraz zbieżny tak szybko jak $O(n^{-1/3})$, a więc nieco wolniej.

UWAGI KOŃCOWE

W pracy identyfikuje się systemy o szeregowej strukturze składające się z nieliniowej części statycznej i liniowej, dynamicznej. Złożoność struktury powoduje, że nieliniowe charakterystyki mogą być wyznaczone jedynie z dokładnością do pewnych stałych. Fakt ten cechuje problem i jest niezależny od proponowanych algorytmów. Należy tytaj podkreślić, że identyfikacja systemów złożonych charakteryzuje się specyficznymi problemami takimi jak duże, na ogół, trudności obliczeniowe oraz ograniczone możliwości wnioskowania spowodowane częstym brakiem możliwości pomiarów niektórych sygnałów systemu. W niniejszej pracy nie są mierzone sygnały łączące podsystemy w systemach Hammersteina i Wienera.

Drugą istotną cechą pracy jest to, że informacja wstępna o badanym systemie jest niezwykle uboga. Jest on tak niewielka, że charakterystyki nieliniowej w żadnym z trzech omawianych zadań nie da się przedstawić w postaci parametrycznej, tzn. nie można podać jej reprezentacji typu $m(u) = \sum_{k=1}^{N} a_k \varphi_k(u)$, gdzie $\varphi_1,...,\varphi_N$ jest znanym układem funkcji, natomiast $a_1,...,a_N$ nieznanymi współczynnikami, które należy estymować. Dlatego też zadania identyfikacji części nieliniowej są nieparametryczne.

Zadania identyfikacji dynamicznych podsystemów liniowych są także nieparametryczne ze względu na to, że rząd ich równań stanu nie jest znany. W niniejszej pracy zagadnienia tego nie poruszaliśmy. Można jednak stwierdzić, że dają się zastosować w tym celu metody korelacyjne. Zainteresowanego czytelnika odsyłamy do prac [1, 2, 5, 6].

LITERATURA

[1] J. S. Bendat, Nonlinear System Analysis and Identification, Wiley, New York, 1990.

[2] S. A. Billings, "Identification of nonlinear systems - a survey", ***Proceedings of IEE***, 127, 1990, 272-285.

[3] G. Giunta, G. Jacovitti, A. Neri, "Bandpass nonlinear system identification by higher cross correlation", ***IEEE Transactions on Signal Processing***, 39, 1991, 2092-2095.

[4] W. Greblicki, "Nieparametryczna identyfikacja systemów", ***Archiwum Automatyki i Robotyki***, 36, 1991, 277-290.

[5] W. Greblicki, "Nonparametric identification of Wiener systems", ***IEEE Transactions on Automatic Control***, 39, 1994, 2077-2086.

[6] W. Greblicki, M. Pawlak, "Nonparametric identification of Hammerstein systems", ***IEEE Transactions on Information Theory***, 35, 1989, 409-418.

[7] B. L. S. Prakasa Rao, Nonparametric Functional Estimation, Academic Press, Orlando, Fl. 1983.