I KONFERENCJA
GRANTY-AUTOMATYKA '95
Teoria sterowania

Warszawa 1995

NIEPARAMETRYCZNA IDENTYFIKACJA SYSTEMÓW †

I KONFERENCJA
GRANTY-AUTOMATYKA '95
Teoria sterowania

Włodzimierz Greblicki‡

Instytut Cybernetyki Technicznej, Wydział Elektroniki, Politechnika Wrocławska
Wybrzeże Wyspiańskiego 27, 50–370 Wrocław

I. WSTĘP

W ramach projektu badawczego opracowano i poddano analizie nieparametryczne algorytmy identyfikacji systemów nieliniowych pracujących w obecności losowych zakłóceń. Algorytmy te posiadają jedną zasadniczą cechę, która odróżnia je od parametrycznych. Mogą one skutecznie identyfikować system w sytuacji, gdy informacja wstępna o nim jest tak mała, że te ostatnie nie mogą być efektywnie zastosowane. O nieznanych charakterystykach nieliniowych zakłada się bowiem jedynie, że są one np. ograniczone lub całkowalne. Klasa wszystkich takich możliwych charakterystyk jest tak bogata, że nie można jej przedstawić w postaci parametrycznej. To, że algorytmy nieparametryczne są skuteczne w tych okolicznościach powoduje, że nabierają one coraz większego znaczenia z punktu widzenia zastosowań. Zaproponowano nieparametryczne algorytmy identyfikacji nieliniowej charakterystyki w systemie statycznym oraz dynamicznym systemie Hammersteina. Wykazano, że są one zbieżne do nieznanych charakterystyk. Oszacowano szybkość tej zbieżności i porównano z tą, którą osiągają algorytmy parametryczne. Przeprowadzono ponadto komputerowe badania symulacyjne.

II. PARAMETRYCZNE I NIEPARAMETRYCZNE ZADANIA IDENTYFIKACJI SYSTEMÓW

Wszelkie, jak się zdaje, wnioskowanie korzysta z dwóch typów informacji. Jedna z nich jest natury empirycznej, zdobywana jest doświadczalnie. Jeśli odnieść się do zagadnień identyfikacji systemów, informację tę stanowią wyniki pomiarów sygnałów w różnych częściach systemu, najczęściej na jego wejściu i wyjściu. Eksperyment może mieć przy tym charakter czynny, lub bierny. Informacja o zupełnie innym charakterze, to ta którą dysponuje się już wcześniej, przed wykonaniem jakiegokolwiek pomiaru czy doświadczenia. Nazywa się ją wstępną, lub aprioryczną. W zadaniach identyfikacji informację tę stanowi nasza wiedza o badanym systemie, którą dysponujemy z góry, już przed przystąpieniem do prowadzenia

† Projekt badawczy KBN Nr 3 P403 001 05. Czas trwania: 1.07.1993–30.06.1994.

‡ Kierownik projektu. Główni wykonawcy: Zygmunt Hasiewicz, Ewaryst Rafajłowicz.

jakichkolwiek pomiarów. Prawidłowe wnioskowanie, czyli w naszym przypadku identyfikacja polega na właściwym wykorzystaniu obydwu typów informacji, informacji o zupełnie różnym charakterze. Jest rzeczą zupełnie oczywistą, że to jaką informacją wstępną dysponujemy w zasadniczy sposób rzutuje na sposoby konstrukcji algorytmów i ich własności.

Jeśli chodzi o informację wstępną, to wyraża się w niej przede wszystkim stopień naszej znajomości badanego systemu. Im więcej o nim wiemy, tym informacja ta jest bogatsza. W zastosowaniach informacja ta jest zazwyczaj skąpa. W literaturze rozważa się dwa zasadnicze rodzaje tej informacji, co zilustrujemy na bardzo prostym przykładzie identyfikacji statycznego systemu nieliniowego o charakterystyce m, tzn. systemu w którym

$$y = m(u),$$

gdzie u jest jego wejściem, a y wyjściem. W literaturze przyjmuje się zwykle, że nasza wiedza wstępna o systemie pozwala stwierdzić, że $m(u) = \sum_{i=1}^{q} \alpha_i \varphi(u)$, przy czym q jest znaną liczbą, a $\{\varphi_i; i = 0,1,2,\dots\}$ ustalonym układem funkcji. Identyfikacja polega na wyznaczeniu, lub estymacji współczynników $\{\alpha_i; i = 0,1,2,\dots\}$ powyższego rozwinięcia. Jeśli identyfikujemy liniowy system dynamiczny opisywany równaniem stanu

$$\begin{aligned} x_{n+1} &= Ax_n + bu_n \\ y_n &= c^T x_n, \end{aligned}$$

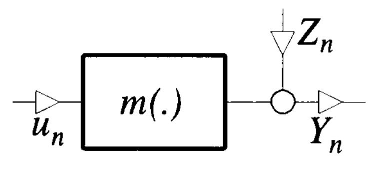
w którym x_n jest wektorem stanu, to przyjmuje się zazwyczaj, że rząd systemu, tzn. stopień macierzy A jest znany. W procesie identyfikacji wyznacza się, lub estymuje wyrazy macierzy A oraz wektorów b i c. O obydwu zadaniach identyfikacji mówimy, że są parametryczne, bowiem nasza niewiedza o nich sprowadza się do nieznajomości skończonej liczby parametrów. Parametry te wyznacza się następnie, lub estymuje w procesie identyfikacji.

Nietrudno o sytuacje, w których nasz informacja wstępna o systemie nie wystarcza do stwierdzenia, że charakterystyka nieliniowa ma podaną powyżej skończoną reprezentację, lub że rząd systemu dynamicznego jest znany. Innymi słowy istnieją sytuacje, w których nasza wiedza wstępna o systemie jest uboższa. Zadania identyfikacji, które należy wtedy rozwiązać nazywamy nieparametrycznymi. W zadaniach takich przyjmuje się jedynie, że nieznana charakterystyka nieliniowa jest np. funkcją ograniczoną, lub całkowalną, lub całkowalną z kwadratem. Zakłada się czasami, że spełnia np. warunek Lipschitza. Każde z tych założeń spełnione jest przez bardzo szerokie klasy funkcji. Klasy te są tak bogate, że nie można ich przedstawić w postaci parametrycznej. Warto tutaj dodać, że niekiedy o charakterystyce tej nie czyni się żadnych założeń.

Jeśli chodzi o system liniowy, to nie nakłada się nań żadnych ograniczeń dotyczących znajomości jego rzędu. Klasa wszystkich wchodzących w rachubę systemów pokrywa się poprostu z klasą wszystkich liniowych systemów dynamicznych. Wynikający stąd problem identyfikacji jest zatem nieparametryczny.

III. SYSTEMY STATYCZNE

Rozpatrzono problem identyfikacji nieliniowej charakterystyki m w sytuacji, gdy $Y_n = m(u_n) + Z_n$, gdzie u_n jest wejściem deterministycznej natury, natomiast Z_n losowym zakłóceniem o charakterze białego szumu z zerową wartością oczekiwaną. W rezultacie wyjście systemu Y_n jest także losowe, patrz Rys. 1, praca [4]. Kolejne wartości $u_1, u_2, \ldots, u_n$ sygnału wejściowego są równomiernie rozłożone na odcinku $[0,1]$. Na podstawie par pomiarów $(u_1, Y_1), \ldots, (u_n, Y_n)$ estymuje się nieznaną charakterystykę m. O charakterystyce tej zakłada się jedynie, że jest całkowalna z kwadratem na przyjętym odcinku. Jest oczywiste, że klasa wszystkich możliwych charakterystyk jest tak bogata, że nie można jej przedstawić w postaci parametrycznej. Zatem problem jej estymacji jest nieparametryczny.



Rys. 1. System statyczny

Zaproponowano następujący algorytm estymacji:

$$m_n(u) = \sum_{k=1}^{q(n)} a_{k,n} \varphi_k(u).$$

W estymatorze tym $\varphi_1(u) = 1$, $\varphi_{2k}(u) = 2^{1/2}\sin(k\pi u)$ oraz $\varphi_{2k+1}(u) = 2^{1/2}\cos(k\pi u)$, natomiast $a_{kn} = \sum_{i=1}^{n} Y_i \varphi(u_i)(u_i - u_{i-1})$ jest estymatorem nieznanego współczynnika $a_k = \int_0^1 m(u)\varphi_k(u)du$ rozwinięcia m w szereg trygonometryczny. Ponadto $\{q(n)\}$ jest odpowiednio wybranym ciągiem liczb całkowitych. Wiadomo, patrz np. [9], że jeśli

$$\lim_{n\to\infty} q(n) = \infty \text{ oraz } \lim_{n\to\infty} q(n)/n = 0,$$

to

$$\lim_{n\to\infty} E\int_{-1}^{1} [m_n(u) - m(u)]^2 du = 0.$$

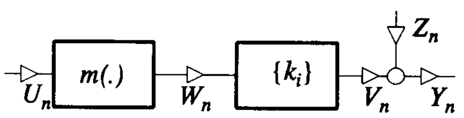
W pracy [4] zbadano numeryczne aspekty podanego estymatora trygonometrycznego. W celu przyspieszenia procedury obliczania jego wartości w zadanych punktach zaproponowano narzędzie jakim jest szybka transformacja Fouriera.

Rozpatrzono także zadania jednoczesnej estymacji charakterystyki i odtwarzania pewnych jej cech jakościowych takich jak monotoniczność lub wypukłość. Założono przy tym, że nie wiadomo z góry, czy cechę taką badana charakterystyka posiada. Algorytm identyfikacji, w którym wykorzystano układ funkcji Bernsteina nie tylko zbiega się do nieznanej charakterystyki, ale także, w probabilistycznym sensie, wykrywa i odtwarza wspomniane wyżej cechy [6].

III. SYSTEMY DYNAMICZNE

3.1. Zadanie identyfikacji

Rozpatrzono problem identyfikacji nieliniowego, dynamiczngo systemu Hammersteina, Rys. 2. W systemie tym statyczny podsystem nieliniowy i liniowy dynamiczny połączone są szeregowo. Sygnał wejściowy jest losowym, białym szumem o zerowej wartości średniej. Jego gęstość prawdopodobieństwa oznaczona jest przez f.



Rys. 2. System Hammersteina

Addytywne zakłócenie pomiarowe o zerowej wartości oczekiwanej może być skorelowane. O nieliniowej charakterystyce m założono jedynie, że

$$Em^2(U_0) < \infty \qquad (3.1)$$

Podsystem liniowy jest asymptotycznie stabilny i posiada nieznaną odpowiedź impulsową $\{k_i; i = 0,1,2,...\}$. Zadanie polega na estymacji nieliniowej charakterystyki m oraz odpowiedzi impulsowej podsystemu dynamicznego. Obydwa problemy są natury nieparametrycznej. Należy tutaj zwrócić uwagę, że, ze względu na szeregową strukturę systemu, zarówno charakterystykę, jak i odpowiedź impulsową można wyznaczyć jedynie z dokładnością do pewnych stałych.

Pewnego komentarza wymaga ograniczenie (3.1). Łącznie z założeniem o asymptotycznej stabilności podsystemu liniowego ma ono na celu jedynie zagwarantowanie, że problem jest poprawnie postawiony. Przez zwrot ten rozumie się tutaj to, że wyjście systemu Y_n jest, dzięki przyjętemu założeniu, zmienną losową. W rezultacie, problem identyfikacji można opisać w języku, którym operuje probabilistyka. Założenie to jest zupełnie niezależne od przyjmowanej metody identyfikacji i nie można uniknąć. Dla przykładu, jest ono spełnione, gdy $EU_0^2 < \infty$ oraz $|m(u)| \le c_1 + c_2 |u|$, dowolne c_1 i c_2.

3.2 Identyfikacja części nieliniowej

3.2.1. Algorytm identyfikacji i jego zbieżność

Punktem wyjściowym dla identyfikacji części nieliniowej jest fakt, że

$$E\{Y_1 \mid U_0 = u\} = \alpha m(u),$$

gdzie α jest pewną stałą niemożliwą ani do wyznaczenia ani estymacji. W celu estymacji charakterystyki nieliniowej, czyli funkcji regresji, zastosowano następujący algorytm:

$$\mu_n(u) = \frac{\sum_{i=1}^{n} Y_{i+1} K\left(\frac{u - U_i}{h(n)}\right)}{\sum_{i=1}^{n} K\left(\frac{u - U_i}{h(n)}\right)}.$$

W algorytmie tym K jest odpowiednio wybraną nieujemną funkcją, a $\{h(n)\}$ pewnych ciągiem dodatnich liczb. Wykazano, że algorytm ten skutecznie wykrywa nieliniowość, co można wyrazić w postaci następującego twierdzenia, patrz praca [1]:

Twierdzenie 1. Załóżmy, że $K(u) = K(-u)$ dla wszystkich u, że $K(u)$ jest nierosnącą funkcją dla dodatnich u, oraz że $\int_{-\infty}^{\infty} K(u)du < \infty$. Załóżmy ponadto, że

$$\lim_{n\to\infty} h(n) = 0 \quad \text{oraz} \quad \lim_{n\to\infty} nh(n) = \infty.$$

Wówczas

$$\mu_n(u) \to \alpha m(u) \quad \text{według prawdopodobieństwa gdy} \quad n \to \infty.$$

Zbieżność ta zachodzi dla prawie wszystkich (według miary Lebesgue'a) punktów u, w których $f(u) > 0$.

Rezultat powyższy dotyczy zbieżności punktowej, czyli własności lokalnej. Jeśli charakterystyka m jest ograniczona, to algorytm zbiega się także w sensie globalnym, o czym stanowi

Twierdzenie 2. Jeśli założenia Twierdzenia 1 są spełnione i ponadto

$$|m(u)| \le c,$$

dowolne c, to

$$\lim_{n\to\infty} E\int_{-\infty}^{\infty} [\mu_n(u) - \alpha m(u)]^2 f(u)du = 0.$$

Istnieje szeroka klasa funkcji, które można przyjąć jako jądro K estymatora, w tym np. następujące: $1/(1+u^2)$, $\exp(-|u|)$, $\exp(-u^2)$. Inny przykład, to jądro prostokątne, równe 1 dla

$|u| \leq 1$ oraz 0 dla pozostałych. Jako ciąg liczbowy $\{h(n)\}$ można przyjąć np. $h(n) \sim n^{-\alpha}$, gdzie $0 < \alpha < 1$.

Trzeba dodać, że podane rezultaty są prawdziwe także w sytuacji, gdy szum pomiarowy jest skorelowany, gdy jest np. wyjściem liniowego systemu dynamicznego pobudzanego białym szumem.

3.2.2. Szybkość zbieżności algorytmu

Szybkość zbieżności algorytmu zależy od stopnia gładkości charakterystyki, mierzonego zazwyczaj liczbą istniejących jej pochodnych. Jest przy tym oczywiste, że na szybkość zbieżności algorytmu rzutuje wybór jądra K i ciągu $\{h(n)\}$. Elementy te należy wybrać w ten sposób, aby szybkość ta była możliwie duża. Dokładniejsza analiza pokazuje, patrz praca [1], że jeśli charakterystyka m i gęstość prawdopodobieństwa sygnału wejściowego f są p-krotnie różniczkowalne oraz $h(n) \sim n^{-p/(2p+1)}$, to

$$|\mu_n(u) - \alpha m(u)| = O(n^{-p/(2p+1)}),$$

według prawdopodobieństwa. W szczególności, jeśli $p = 2$, to otrzymana szybkość zbieżności wynosi $O(n^{-2/5})$. Warto zwrócić uwagę no to, że gdy liczba istniejących pochodnych rośnie do nieskończoności, to podana szybkość staje się bardzo bliska $n^{-1/2}$, tzn. szybkości typowej dla wnioskowania parametrycznego. Otrzymany rezultat należy więc uznać za zachęcający.

Badania symulacyjne sugerują, że własności algorytmu można poprawić, tzn. można zwiększyć szybkość zbieżności, jeśli $h(n)$ uzależni się ponadto od wyników eksperymentu. Przedstawiono odpowiednie propozycje i poddano je weryfikacji numerycznej, [1].

3.2.3. Inne algorytmy

Przedstawiony algorytm identyfikacjiposiada pewną wadę charakterystyczną dla wszystkich znanych algorytmów. Jego szybkość zbieżności jest wolna, gdy gęstość prawdopodobieństwa sygnału wejściowego jest nieregularna, np. gdy jest ona funkcją nieciągłą. Zaproponowany w pracy [3] algorytm wady tej nie posiada. Szybkość jego zbieżności nie zależy bowiem od kształtu tej gęstości. Rezultat ten osiągnięto dzięki temu, że ciąg par obserwacji $(U_1,Y_1),(U_2,Y_2),\dots,(U_n,Y_n)$ zamieniono na nowy $(U_{(1)},Y_{[1]},(U_{(2)},Y_{[2]}),\dots,(U_{(n)},Y_{[n]})$, w którym pary uporządkowano w ten sposób, że $U_{(1)} < U_{(2)} < \dots < U_{(n)}$. Algorytm skonstruowano następnie otrzymanego w ten sposób ciągu. Jego całkowy błąd średniokwadratowy zbiega się do zera z szybkością równa $O(n^{-2p/(2p+1)})$, gdzie p jest liczbą istniejących pochodnych estymowanej charakterystyki.

Jeśli sygnał wejściowy jest skorelowany, to zadanie komplikuje się znacznie. Dla estymacji m należy bowiem rozwiązać zadanie dekonwolucji. Propozycję konstrukcji algorytmu identyfikacji skutecznego w takiej sytuacji przedstawiono w pracy [2].

3.3. Identyfikacja części dynamicznej

Zaproponowany algorytm identyfikacji liniowej części dynamicznej, tzn. estymacji jego odpowiedzi impulsowej wychodzi z obserwacji, że

$$\mathrm{cov}(Y_i, U_0) = \beta k_i,$$

gdzie $\beta = \mathrm{cov}(m(U_0), U_0)$. Dla estymacji βk_i zaproponowano następującą procedurę:

$$\kappa_{i,n} = n^{-1} \sum_{j=1}^{n} Y_{i+j} U_j.$$

Wykazano, że jej zbieżność, praca [1].

IV. INNE WYNIKI I ZAKOŃCZENIE

Opracowano ponadto zagadnienia identyfikacji parametrycznej, które pozostają w bezpośrednim związku z tematem projektu. Podano efektywny numerycznie algorytm rozwiązywania zadań związanych z metodą najmniejszych kwadratów [5]. Wyniki będzie można wykorzystać w przyszłości w identyfikacji nieparametrycznej, gdyż znane są sposoby adaptacji metody najmniejszych kwadratów dla zadań nieparametrycznych. Prowadzono badania porównawcze nad parametryczną i nieparametryczną identyfikacją systemów statycznych z złożonej strukturze [7,8]. W szczególności podano algorytm optymalnej aproksymacji w sytuacji nieparametrycznej.

W ramach projektu poczyniono dalszy krok w zakresie rozwoju nieparametrycznych metod identyfikacji systemów.

WYKAZ LITERATURY

Wykaz ważniejszych prac wykonanych w ramach projektu

[1]. W. Greblicki, M. Pawlak, "Casacade nonlinear system identification by nonparametric method", ***International Journal of Systems Science***, vol. 25, str. 129–153, 1994.

[2]. W. Greblicki, M. Pawlak, "Nonlinear system identification with nonparametric deconvolution", ***Proceedings IEEE International Symposium on Information Theory***, 27 czerwca – 1 lipca 1994, Trondheim, Norwegia, str. 124, 1994.

[3]. M. Pawlak, W. Greblicki, "Nonlinear system identification with the help of order statistics" ***Proceedings 10th IFAC Conference on System Identification***, 4 – 6 lipca, 1994, Kopenhaga, Dania, str. 151–161, 1994.

[4]. E. Rafajłowicz, E. Skubalska–Rafajłowicz, "FFT in calculating nonparametric regression estimate based on trigonometric series", ***Applied Mathematics and Computer Science***, vol. 3, str. 713-720, 1993.

[5]. E. Rafajłowicz, W. Myszka, "Efficient algorithm for a class of least squares estimation problems", ***IEEE Transactions on Automatic Control***, vol. 39, 1994.

[6]. E. Rafajłowicz, "Nonparametric shape preserving regression function estimation using Bernstein-Durrmeyer polynomials", (referat wygłoszony w Freie Universität, Berlin West, Niemcy), ***Raport Instytutu Cybernetyki Technicznej***, Nr 50/93, str. 14, 1993.

[7]. Z. Hasiewicz, "Identyfikacja sterowanych systemów o złożonej strukturze", ***Prace Naukowe Instytutu Cybernetyki Technicznej Politechniki Wrocławskiej***, Nr 90, (rozprawa habilitacyjna), str. 190, 1994.

[8]. Z. Hasiewicz, "Two-stage algorithm for the linear approximations of large interconnected steady-state systems", ***International Journal of Systems Science***, vol. 23, 1994.

Inne prace

[9] E. Rafajłowicz, "Nonparametric orthogonal series estimators of regression: A class attaining the optimal convergence rate in L_2", ***Statistics and Probability Letters***, vol. 5, str. 219-224, 1987.