PODSTAWY STEROWANIA
Tom 4 (1974), z. 1
pp. 3—12

Identyfikacja statyczna metodą szeregów ortogonalnych

WŁODZIMIERZ GREBLICKI

(Maszynopis wpłynął 28 października 1973)

1. Wstęp

Identyfikacja obiektu statycznego w warunkach probabilistycznych polega z reguły na odpowiedniej aproksymacji przy pomocy liniowego układu funkcji ortogonalnych. Wyniki obserwacji wejścia i wyjścia służą do estymacji współczynników modelu najlepszego w tak określonej klasie [3, 6]. Z uwagi na to, że aproksymuje się przy pomocy szeregu skończonego, dokładność identyfikacji jest z góry ograniczona. Można ją oczywiście polepszyć wybierając odpowiednio dużą ilość funkcji w kombinacji liniowej określającej model. W szczególności można nie ograniczać klasy dopuszczalnych rozwiązań i identyfikować posługując się modelem opisanym liniową kombinacją nieograniczonej ilości funkcji. Pojawiający się w ten sposób nieparametryczny problem identyfikacji można rozwiązać stosując algorytm metody funkcji potencjalnych [1, 2]. Zbieżność procesu identyfikacji do funkcji regresji charakteryzującej najlepszy w sensie kwadratowym, model zapewnia także nieparametryczna metoda podana w pracy [4].

W poniższej pracy przedstawiono nieparametryczny algorytm identyfikacji metodą rozwinięcia funkcji regresji w szereg ortogonalny i wykazano jego zbieżność.

2. Algorytm identyfikacji

Załóżmy, że obiekt o skalarnym wejściu u i wyjściu y identyfikuje się w sytuacji, gdy znana jest gęstość $f(u)$ sygnału wejściowego, natomiast całkowicie nieznana jest gęstość warunkowa $f(y/u)$ charakteryzująca własności obiektu. Obserwacje

$$(u_1, y_1), (u_2, y_2), \dots, (u_n, y_n)$$

wejścia i wyjścia są realizacjami niezależnych zmiennych losowych

$$(U_1, Y_1), (U_2, Y_2), \ldots, (U_n, Y_n)$$

o łącznej gęstości

$$f(u, y) = f(y/u) f(u) .$$

Zmienne Y_i przyjmują wartości z przestrzeni $\mathcal{Y}$ wszystkich liczb rzeczywistych, natomiast U_i z podzbioru $\mathcal{U}$ tej przestrzeni. Oznaczmy jeszcze przez (U, Y) parę zmiennych losowych o rozkładzie $f(u, y)$.

Załóżmy teraz, że

$$EY^2 < \infty , \tag{1}$$

$$\sup_u f(u) < \infty \tag{2}$$

i oznaczmy

$$\Phi(u) \stackrel{\mathrm{df}}{=} \int_{\mathcal{Y}} y f(y/u)\, dy , \tag{3}$$

$$R(u) \stackrel{\mathrm{df}}{=} \int_{\mathcal{Y}} y f(u, y)\, dy = f(u) \Phi(u) . \tag{4}$$

Niech $\{\varphi_j(u)\}_{j=0}^{\infty}$ będzie pełnym układem funkcji ortonormalnych w przestrzeni L_2 funkcji całkowalnych z kwadratem na zbiorze $\mathcal{U}$. Załóżmy ponadto, że funkcje tego układu są wspólnie ograniczone, tzn. że

$$\sup_{j,u} |\varphi_j(u)| \leqslant c < \infty . \tag{5}$$

Z założeń (1) i (2) wynika, że $R(u) \in L_2$, ponieważ

$$\int_{\mathcal{U}} R^2(u)\, du = \int_{\mathcal{U}} f^2(u) \Big[\int_{\mathcal{Y}} y f(y/u)\, dy \Big]^2 du \leqslant \int_{\mathcal{U}} \int_{\mathcal{Y}} y^2 f(u, y) f(u)\, du\, dy \leqslant$$
$$\leqslant [\sup_u f(u)] EY^2 < \infty .$$

Funkcję $R(u)$ można więc rozłożyć w szereg

$$\sum_{j=0}^{\infty} a_j \varphi_j(u)$$

o współczynnikach określonych wzorami

$$a_j = \int_{\mathcal{U}} R(u) \varphi_j(u)\, du = E\{Y \varphi_j(U)\} .$$

Będziemy je estymować w naturalny sposób, a mianowicie

$$a_{jn} = \frac{1}{n} \sum_{i=1}^{n} y_i \varphi_j(u_i) . \tag{6}$$

Estymator $R_n(u)$ nieznanej funkcji $R(u)$ określimy wzorem

$$R_n(u) = \sum_{j=0}^{N} a_{jn}\varphi_j(u),$$

gdzie N jest odpowiednio, w zależności od n, ustalaną liczbą, a estymator $\Phi_n(u)$ funkcji regresji $\Phi(u) = E(y/u)$ zdefiniujemy wzorem

$$\Phi_n(u) = \frac{1}{f(u)} \sum_{j=0}^{N} a_{jn}\varphi_j(u). \tag{7}$$

3. Zbieżność algorytmu

Jest oczywiste, że

$$Ea_{jn} = a_j. \tag{8}$$

Stąd i z założenia (5) wynika, że

$$E\{a_j - a_{jn}\}^2 = E\left\{\frac{1}{n}\sum_{i=1}^{n} Y_i\varphi_j(U_i)\right\}^2 - a_j^2 = \frac{1}{n}(E\{Y^2\varphi_j^2(U)\} - a_j^2) \leqslant$$

$$\leqslant \frac{1}{n}c^2EY^2 = \frac{c_1}{n}. \tag{9}$$

Wykorzystując równość Parsevala otrzymujemy z kolei

$$E\int_u [R(u) - R_n(u)]^2 du = \sum_{j=0}^{N} E\{a_j - a_{jn}\}^2 + \sum_{j=N+1}^{\infty} a_j^2 \leqslant c_1\frac{N}{n} + \sum_{j=N+1}^{\infty} a_j^2.$$

Jeśli teraz N uzależnić od ilości obserwacji n w ten sposób, aby

$$\lim_{n\to\infty} N(n) = \infty, \quad \lim_{n\to\infty}\frac{N(n)}{n} = 0, \tag{10}$$

to z powyższej nierówności wynika, że

$$\lim_{n\to\infty} E\int_u [R(u) - R_n(u)]^2 du = 0, \tag{11}$$

tzn.

$$\lim_{n\to\infty} E\int_u [\Phi(u) - \Phi_n(u)]^2 f^2(u) du = 0.$$

Otrzymany wynik podamy teraz w postaci twierdzenia.

TWIERDZENIE

Jeśli $\{\varphi_j(u)\}_{j=0}^{\infty}$ jest pełnym układem ortonormalnym w przestrzeni L_2 takim, że

$$\sup_{j,u} |\varphi_j(u)| < \infty; \quad EY^2 < \infty$$

oraz

$$\lim_{n\to\infty} N(n) = \infty, \quad \lim_{n\to\infty} \frac{N(n)}{n} = 0,$$

to

$$\lim_{n\to\infty} E \int_{u} [\Phi(u) - \Phi_n(u)]^2 f^2(u)\, du = 0, \tag{12}$$

gdzie

$$\Phi_n(u) = \frac{1}{f_n(u)} \sum_{j=0}^{N(n)} a_{jn} \varphi_j(u)$$

oraz

$$a_{jn} = \frac{1}{n} \sum_{i=1}^{n} Y_i \varphi_j(U_i).$$

Podamy teraz kilka przykładowych algorytmów identyfikacji.

Przykład 1

Niech $\mathcal{U}$ będzie prostą $(-\infty, +\infty)$, a $\{\varphi_j(u)\}$ ortonormalnym układem Hermite'a, tzn.

$$\varphi_j(u) = (2^j j!\, \pi^{1/2})^{-1/2} e^{-\frac{u^2}{2}} H_j(u), \quad j = 0, 1, \ldots,$$

przy czym

$$H_j(u) = (-1)^j e^{u^2} \frac{d^j}{du^j} e^{-u^2}$$

jest j-ym wielomianem Hermite'a. Układ ten, jak wiadomo [7], spełnia założenie (5).

Przykład 2

Niech $\mathcal{U}$ będzie odcinkiem $[0, 2\pi]$. Jako $\{\varphi_j(u)\}$ można wybrać układ trygonometryczny

$$\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\cos u, \frac{1}{\sqrt{\pi}}\sin u, \frac{1}{\sqrt{\pi}}\cos 2u, \frac{1}{\sqrt{\pi}}\sin 2u, \ldots$$

Niech dalej

$$\frac{a_0}{\sqrt{2\pi}} + \sum_{k=1}^{\infty} \frac{1}{\sqrt{\pi}} (a_k \cos ku + \beta_k \sin ku)$$

będzie rozwinięciem funkcji $R(u)$. Współczynniki a_k i β_k estymuje się następująco

$$a_{0n} = \frac{1}{n\sqrt{2\pi}} \sum_{i=1}^{n} y_i\,, \qquad a_{kn} = \frac{1}{n\sqrt{\pi}} \sum_{i=1}^{n} y_i \cos ku_i\,, \qquad \beta_{kn} =$$

$$= \frac{1}{n\sqrt{\pi}} \sum_{i=1}^{n} y_i \sin ku_i\,, \qquad k = 1, 2, \ldots$$

Estymator funkcji $R(u)$ wyraża się więc wzorem

$$R_n(u) = \frac{a_{0n}}{\sqrt{2\pi}} + \sum_{k=1}^{N(n)/2} \frac{1}{\sqrt{\pi}} (a_{kn} \cos ku + \beta_{kn} \sin ku)\,.$$

Jeśli teraz ciąg $N(n)$ spełnia założenia twierdzenia, to jest oczywiste, że zachodzi (11), a zatem i (12), przy czym

$$\Phi_n(u) = \frac{1}{f(u)} \left[\frac{a_{0n}}{\sqrt{2\pi}} + \sum_{k=i}^{N(n)/2} \frac{1}{\sqrt{\pi}} (a_{kn} \cos ku + \beta_{kn} \sin ku) \right].$$

4. Zgodność estymatora funkcji regresji

Wybierając odpowiednio układ $\{\varphi_j(u)\}$ i ciąg $N(n)$ można otrzymać zgodny estymator funkcji regresji. Oznaczmy w tym celu

$$R^N(u) \stackrel{\mathrm{df}}{=} \sum_{j=0}^{N} a_j \varphi_j(u)\,.$$

Zatem, dla $N = N(n)$

$$E\{R(u) - R_n(u)\}^2 = E\{R(u) - R^N(u) + R^N(u) - R_n(u)\}^2 =$$
$$= [R(u) - R^N(u)]^2 + E\{R^N(u) - R_n(u)\}^2 + (R(u) - R^N(u)) E\{R^N(u) - R_n(u)\} =$$
$$= [R(u) - R^N(u)]^2 + E\{R^N(u) - R_n(u)\}^2\,,$$

onieważ, jak wynika z (8), $ER_n(u) = R^N(u)$. Zauważmy teraz, że na podstaie (9) i nierówności Schwartza mamy

$$E\{R^N(u) - R_n(u)\}^2 = E\left\{ \sum_{j=0}^{N} (a_j - a_{jn}) \varphi_j(u) \right\}^2 = \sum_{j=0}^{N} E\{a_j - a_{jn}\}^2 \varphi_j^2(u) +$$
$$+ \sum_{\substack{k,l \\ k \neq l}} E\{(a_k - a_{kn})(a_l - a_{ln})\} \varphi_k(u) \varphi_l(u) \leqslant c^2 c_1^2 \frac{N^2}{n} \stackrel{\mathrm{df}}{=} c_2 \frac{N^2}{n}\,.$$

Ostatecznie

$$E\{R(u)-R_n(u)\}^2 \leqslant \Big[\sum_{j=0}^{N} a_j\varphi_j(u)-R(u)\Big]^2 + c_2\frac{N^2}{n}. \tag{13}$$

Jeśli teraz układ ortonormalny $\{\varphi_j(u)\}$ i ciąg $N(n)$ wybierze się tak, aby spełnione były warunki

$$\lim_{n\to\infty} N(n) = \infty , \quad \lim_{n\to\infty}\frac{N^2(n)}{n} = 0 \tag{14}$$

oraz

$$\lim_{N\to\infty}\Big[R(u) - \sum_{j=1}^{N} a_j\varphi_j(u)\Big]^2 = 0 \tag{15}$$

w ustalonym punkcie $u \in \mathfrak{U}$, to z nierówności (13) wynika, że

$$\lim_{n\to\infty} E\{R(u)-R_n(u)\}^2 = 0 , \tag{16}$$

a zatem

$$\lim_{n\to\infty} E\{\Phi(u)-\Phi_n(u)\}^2 = 0 \tag{17}$$

w tymże punkcie $u \in \mathfrak{U}$, co oznacza, że estymator (7) jest zgodny.

Poniżej podano dwa przykłady właściwego wyboru układu $\{\varphi_j(u)\}$ i ciągu $N(n)$.

Przykład 3

Niech, podobnie jak w przykładzie 1, $\mathfrak{U}$ będzie całą prostą, a $\{\varphi_j(u)\}$ ortonormalnym układem Hermite'a. Zauważmy, że z założenia (1) wynika, że $R(u) \in L_1$ oraz, co już wykazano, $R(u) \in L_2$. Jeśli ponadto $R(u)$ jest funkcją ciągłą o ograniczonym wahaniu, to zbieżność (15), a zatem i (17) ma miejsce dla wszystkich $u \in \mathfrak{U}$ [7].

Przykład 4

Niech $U = [0, 2\pi]$ i $\{\varphi_j(u)\}$ będzie układem trygonometrycznym. Jeśli ciąg $N(n)$ ma tę własność, że

$$\frac{N(n+1)}{N(n)} \geqslant \varkappa > 1 , \tag{18}$$

to [8] (oznaczenia, jak w przykładzie 1)

$$\lim_{n\to\infty}\Big[\frac{a_0}{\sqrt{2\pi}} + \sum_{k=1}^{N(n)/2}\frac{1}{\sqrt{\pi}}(a_k\cos ku+\beta_k\sin ku)\Big] = R(u)$$

prawie wszędzie. Jest oczywiste, że własność ta jest zachowana także w sytuacji, gdy dla pewnych n nie jest spełniony warunek (18), a ciąg $N(n)$ ma powtórzenie. Ciąg ten musi wówczas spełniać oczywisty warunek

$$\lim_{n\to\infty} N(n) = \infty .$$

Przykładem może być ciąg o odpowiednio powtarzających się elementach $1, 2, 3, 5, 8, \ldots$ ($\varkappa = 1, 5$). Ilości powtórzeń należy dobrać w ten sposób, aby spełniony został drugi z warunków (14).

5. Identyfikacja przy nieznanym rozkładzie wejścia

Podamy teraz zgodny algorytm identyfikacji w sytuacji, gdy gęstość $f(u)$ sygnału wejściowego nie jest znana. Jest oczywiste, że jeśli w pewnym punkcie $u \in \mathfrak{U}$

$$R_n(u) \to R(u) \text{ według prawdopodobieństwa}$$

gdy $n \to \infty$, $f_n(u)$ jest zgodnym estymatorem gęstości, tzn.

$$f_n(u) \to f(u) \text{ według prawdopodobieństwa}$$

gdy $n \to \infty$ i ponadto $f(u) > 0$, to

$$\Phi_n'(u) \overset{\mathrm{df}}{=} \frac{R_n(u)}{f_n(u)}$$

jest zgodnym estymatorem funkcji regresji.

Jako $R_n(u)$ można zastosować algorytmy podane w przykładach 3 i 4, natomiast $f_n(u)$ może być dowolnym, nieparametrycznym i zgodnym estymatorem gęstości prawdopodobieństwa [5, 7, 9]. Poniżej podano dla przykładu estymatory, które otrzymuje się przez rozwinięcie gęstości w szereg ortogonalny.

Niech $\mathfrak{U}$ będzie prostą. Jeśli $f(u)$ jest funkcją ciągłą o ograniczonym wahaniu i $f(u) \in L_2$ oraz

$$f_n(u) = \sum_{j=0}^{N(n)} b_{jn} \varphi_j(u) , \tag{19}$$

gdzie $\{\varphi_j(u)\}$ jest układem Hermite'a, przy czym

$$b_{jn} = \frac{1}{n} \sum_{i=1}^{n} \varphi_j(u_i) \tag{20}$$

i ponadto, podobnie jak w przykładzie 3, ciąg $N(n)$ spełnia warunki (14), to $f_n(u)$ jest zgodnym estymatorem gęstości $f(u)$ w każdym punkcie $u \in \mathfrak{U}$ [7].

Algorytm identyfikacji funkcji regresji ma więc postać

$$\Phi'_n(u) = \frac{\sum_{j=0}^{N(n)} a_{jn}\varphi_j(u)}{\sum_{j=0}^{N(n)} b_{jn}\varphi_j(u)} . \tag{21}$$

Załóżmy teraz, że $\mathcal{U} = [0, 2\pi]$ i $\{\varphi_j(u)\}$ jest układem trygonometrycznym, podobnie jak w przykładzie 4. Jak wiadomo, jeśli estymator gęstości ma postać (19) oraz ciąg $N(n)$ spełnia warunki (14), to [5]

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

Wykorzystując własności rozwinięcia w szereg trygonometryczny (podobnie jak w przykładzie 4), można łatwo wykazać, że jeśli ciąg $N(n)$ wybierze się jak w przykładzie 4, to estymator (19) jest zgodny prawie wszędzie. Zatem i algorytm identyfikacji (21) jest również zgodny prawie wszędzie.

6. Uwagi końcowe

W algorytmach podanych w pracy zakładano, że wejście obiektu jest skalarne, ale wyniki można łatwo rozszerzyć na przypadki, w których u jest wektorem. Jako układ ortonormalny można wówczas przyjąć układ wszystkich funkcji typu

$$\varphi_{j_1}(u^{(1)})\varphi_{j_2}(u^{(2)})\ldots\varphi_{j_p}(u^{(p)}) ,$$

gdzie p jest wymiarem wejścia, a $u^{(i)}$ składową wektora u.

Static Identification by Orthogonal Seriés Method

A class of nonparametric estimators of the regression curve $E(y/u)$ is considered. The identified object is characterized by an absolutely unknown conditional density $f(y/u)$. (u_i, y_i), $i = 1, 2, \ldots, n$, are observed input and output values.

Let $\{\varphi_j(u)\}_{j=0}^{\infty}$ be a complete orthonormal set over a set $\mathcal{U}$. It is proved that under some additional conditions

$$\lim_{n\to\infty} E \int_{u} [E(y/u) - \Phi_n(u)]^2 f^2(u)\, du = 0$$

($f(u)$ is the input probability density) where

$$\Phi_n(u) = \frac{1}{f(u)} \sum_{j=0}^{N(n)} a_{jn}\varphi_j(u)$$

and

$$a_{j_n} = \frac{1}{n} \sum_{i=1}^{n} y_i\varphi_j(u_i)$$

and

$$\lim_{n\to\infty} N(n) = \infty\,, \quad \lim_{n\to\infty} \frac{N(n)}{n} = 0\,.$$

In the case when $f(u)$ is unknown an estimator

$$\frac{\sum_{=0}^{N(n)} a_{jn}\varphi_j(u)}{\sum_{j=0}^{N(n)} b_{jn}\varphi_j(u)}$$

where

$$b_{jn} = \frac{1}{n}\sum_{i=1}^{n} \varphi_j(u_i)$$

is proposed and its consistency is considered.

Examples of estimators which can be obtained using the Hermite and trigonometric orthonormal sets are given.

Статическая идентификация методом ортогональных рядов

В работе представляется класс непараметрических оценок функции регрессии. Объект о входе u и выходе y описывается плотностью вероятности $f(y/u)$. Пусть $\{\varphi_j(u)\}_{j=0}^{\infty}$ полная система ортонормальных функций на множестве $\mathfrak{U}$. Дается теорема:
если

$$\lim_{n\to\infty} N(n) = \infty\,, \quad \lim_{n\to\infty} \frac{N(n)}{n} = \infty\,,$$

то, при дополнительных предположениях,

$$\lim_{n\to\infty} E \int_{u} [E(y/u) - \Phi_n(u)]^2 f^2(u)\, du\,,$$

где

$$\Phi_n(u) = \frac{1}{f(u)} \sum_{j=0}^{N(n)} a_{jn}\varphi_j(u)$$

и

$$a_{jn} = \frac{1}{n}\sum_{i=1}^{n} y_i\varphi_j(u_i)\,.$$

В случае, когда $f(u)$ неизвестна представляется оценка

$$\frac{\sum_{j=0}^{N(n)} a_{jn}\varphi_j(u)}{\sum_{j=0}^{N(n)} b_{jn}\varphi_j(u)}\,, \quad b_{jn} = \frac{1}{n}\sum_{i=1}^{n} \varphi_j(u_i)\,.$$

Рассматривается её состоятельность.

Представляются примеры оценок, которые можно получить для ортонормальной системы Гермита и системы тригонометрической.

Literatura

[1] M. A. Aizerman, E. M. Brawerman, L. I. Rozonoer, *Metod potencialnych funkcij w zadacze o wosstanowlenii charakteristiki funkcionalnogo preobrazowatelia po słuczajno nabludajemym toczkam*, Awtomatika i Telemechanika, t. XXV, nr 12, 1964.

[2] E. M. Brawerman, *O metodie potencialnych funkcij*, Awtomatika i Telemechanika, t. XXVI, nr 12, 1965.

[3] Z. Bubnicki, *System Identification via Estimation of Probability Distribution and Moments*, 2nd IFAC Symposium on Identification, Prague 1970.

[4] W. Greblicki, *Nieparametryczna identyfikacja charakterystyk statycznych*, Podstawy Sterowania, t. 3, nr 1, 1973.

[5] R. Kronmal, M. Tarter, *The Estimation of Probability Densities and Cumulatives by Fourier Series Methods*, J. Amer. Statist. Assoc., vol. 63, 1968.

[6] K. Mańczak, *Metody identyfikacji wielowymiarowych obiektów sterowania*, WNT, Warszawa 1971.

[7] S. C. Schwartz, *Estimation of Probability Density by an Orthogonal Series*, Ann. Math. Statist., vol. 38, 1967.

[8] R. Sikorski, *Funkcje rzeczywiste*, PWN, Warszawa 1958.

[9] T. Cacoullos, *Estimation of a Multivariate Density*, Ann. Inst. Statist. Math., vol. 18, 1966.

DR INŻ. W. GREBLICKI

INSTYTUT CYBERNETYKI TECHNICZNEJ POLITECHN. WROCŁ.

ZAKŁAD SYSTEMÓW STEROWANIA

JANISZEWSKIEGO 11/17, 50-372 WROCŁAW, POLSKA