Włodzimierz Greblicki
Politechnika Wrocławska

Prace V Krajowej Konferencji Automatyki
Gdańsk, 7 - 9 czerwca 1971
zeszyt 1, str 19-24

1.3.

NIEPARAMETRYCZNA IDENTYFIKACJA OBIEKTÓW STATYCZNYCH

1. Wstęp

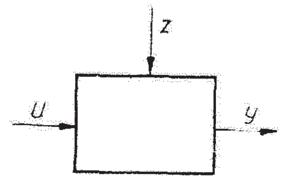
Znane metody rozwiązania probabilistycznego problemu identyfikacji stacjonarnego obiektu statycznego polegają na przyjęciu klasy charakterystyk modeli, w której następnie minimalizuje się wskaźnik jakości indentyfikacji. Niekiedy nieznany, łączny rozkład prawdopodobieństwa wielkości wejściowych i wyjściowych aproksymuje się funkcjami z wybranej klasy i otrzymane przybliżenie traktuje jako rozkład rzeczywisty. Obydwa sposoby mogą, na skutek zawężenia rozważań tylko do pewnej klasy funkcji, prowadzić do niedokładnych wyników. Przedstawiona w poniższej pracy nieparametryczna metoda indentyfikacji pozwala uniknąć tych błędów. Wyznacza ona model najlepszy spośród wszystkich modeli.

2. Założenia

Przez u będzie oznaczane wejście, a przez y wyjście badanego obiektu o schemacie blokowym jak na rys. 1. Wektor wejściowy u należy do k-wymiarowej przestrzeni U, a skalar y do przestrzeni Y liczb rzeczywistych. Na skutek działania losowych zakłóceń z wyjście, nawet przy ustalonych wartościach wejścia, zmieniałoby się przypadkowo. Podczas eksperymentu wejście, a zatem i wyjście, zmieniają się losowo. Obserwacje wykonywane są w kolejnych chwilach. Wyniki pomiarów

$$u_1, y_1; u_2, y_2; \dots u_n, y_n \qquad (1)$$

są realizacjami odpowiednich zmiennych losowych, których łączna gęstość prawdopodobieństwa $f(u,y)$ nie zmienia się w czasie. Wyniki obserwacji wykonywanych w różnych chwilach są niezależne. Gęstość $f(u,y)$ jest nieznana i jedyną informacją o obiekcie są pomiary (1).



Rys. 1 Schemat blokowy badanego obiektu

3. Problem identyfikacji

Zadanie identyfikacji polega na znalezienie charakterystyki $\phi^*(u)$, która minimalizuje wskaźnik jakości

$$\overline{Q}[\phi(u)] = \int\int Q(y-\phi(u)) f(u,y)\, du\, dy, \qquad (2)$$

przy czym $Q(\cdot)$ jest funkcją strat oceniającą różnicę pomiędzy wyjściami obiektu i modelu, który posiada charakterystykę $\phi(u)$

Powyższy problem nieparametryczny można uprościć ograniczając charakterystyki modeli do pewnej, dość arbitralnie wybranej, rodziny funkcji $\phi(u,a)$. Wybór modelu polega teraz na ustaleniu wektora a . Zadanie identyfikacji sprowadza się do znalezienia takiego a^* , które minimalizuje wyrażenie

$$\bar{Q}[\phi(u,a)] = \int_u \int_y Q\big(y - \phi(u,a)\big) f(u,y)\, du\, dy. \tag{3}$$

Innym uproszczeniem zadania (1) jest aproksymacja nieznanej gęstości $f(u,y)$ funkcją z pewnej klasy $\bar{f}(u,y,b)$ wybranej przez eksperymentatora [2],[4] . Wektor b ustala się minimalizując wskaźnik

$$Q_{apr}[f(u,y), \bar{f}(u,y,b)], \tag{4}$$

który ocenia jakość aproksymacji. Charakterystykę modelu $\phi_1(u,b^*)$ wyznacza się poprzez minimalizację względem $\phi(u)$ wyrażenia

$$\int_u \int_y Q\big(y - \phi(u)\big) \bar{f}(u,y,b^*)\, du\, dy, \tag{5}$$

gdzie b^* jest wektorem najlepszej, w sensie (4) , aproksymacji nieznanego rozkładu $f(u,y)$.

Najdalej idącym uproszczeniem jest jednoczesne ograniczenie klasy modeli do rodziny funkcji $\phi(u,a)$ i aproksymacja nieznanego rozkładu [2],[4] . Dla wyznaczenia modelu należy rozwiązać zadanie minimalizacji

$$\min_a \int_u \int_y Q\big(y - \phi(u,a)\big) f(u,y,b^*)\, du\, dy \tag{6}$$

Oznaczmy przez $\phi(u,a^*,b^*)$ otrzymaną w ten sposób charakterystykę modelu.

Wzajemne powiązania powyższych czterech zadań pokazano w formie graficznej (strzałki oznaczają kolejne stopnie uproszczenia problemu) .

I. Problem nieparametryczny

II. Klasa modeli określona parametrycznie.

III. Parametryczna aproksymacja gęstości.

IV. Klasa modeli określona parametrycznie. Parametryczna aproksymacja gęstości.

Z oczywistych nierówności

$$\bar{Q}[\phi^*(u)] \leqslant \bar{Q}[\phi(u,a^*)] \leqslant \bar{Q}[\phi(u,a^*,b^*)],$$
$$\bar{Q}[\phi^*(u)] \leqslant \bar{Q}[\phi_1(u,b^*)] \leqslant Q[\phi(u,a^*,b^*)].$$

wynika, że kolejne uproszczenia wyjściowego problemu I prowadzą do pogorszenia rezultatów identyfikacji. Strzałki na wykresie oznaczają więc również wzrost wartości wskaźnika.

Jakości rozwiązań problemów II i III zależą od sposobu przyjęcia klas $\phi(u,a)$ i $\bar{f}(u,y,b)$ i dlatego nie można z góry rozstrzygnąć, dla którego z nich kryterium (2) przyjmie mniejszą wartość. Pewną wadą w sposobie sformułowania problemu III jest brak powiązania pomiędzy oceną jakości aproksymacji gęstości (4), a zasadniczym kryterium jakości identyfikacji.

Sposoby rozwiązania zagadnień II, III i IV są ogólnie znane [1]-[3]. W tej pracy podano metodę rozwiązania nieparametrycznego problemu I. Pozwala ona wyznaczyć ciąg charakterystyk $\phi_n(u)$ zbieżnych do najlepszej $\phi^*(u)$.

4. Rozwiązanie problemu nieparametrycznego

Informację o obiekcie zdobytą podczas eksperymentu można wykorzystać do konstrukcji empirycznych gęstości prawdopodobieństwa według następującego wzoru [6],[7]:

$$f_n(u,y) \overset{df}{=} \frac{1}{nh_n^{k+1}} \sum_{i=1}^{n} K\left(\frac{u-u_i}{h_n}, \frac{y-y_i}{h_n}\right). \tag{7}$$

Jeśli gęstość $f(u,y)$ jest ograniczona oraz funkcja $k(u,y)$ spełnia warunki

$$0 \leqslant K(u,y) < \infty, \int_u \int_y K(u,y)\,du\,dy, \int_u \int_y K^2(u,y)\,du\,dy \tag{8}$$

i

$$h_n > 0, \qquad h_n \longrightarrow 0, \qquad nh_n^{2K+1} \longrightarrow \infty \tag{9}$$

gdy $n \to \infty$, to [5]

$$\lim_{n\to\infty} E \int_u \int_y [f_n(u,y) - f(u,y)]^2\,du\,dy = 0.$$

Podany ciąg gęstości empirycznych jest więc zbieżny do gęstości rzeczywistej.

Charakterystykę modelu $\phi_n(u)$ wyznacza się minimalizując

$$\int_u \int_y Q(y-\phi(u)) f_n(u,y)\,du\,dy$$

lub, co na jedno wychodzi,

$$\int_y Q(y-\phi(u)) f_n(u,y)\,du\,dy.$$

Po uwzględnieniu wzoru (7) powyższe wyrażenie przyjmuje postać

$$\sum_{i=1}^{n} \int_y Q(y-\phi(u)) K\left(\frac{u-u_i}{h_n}, \frac{y-y_i}{h_n}\right) dy. \tag{10}$$

Jeśli ... funkcja $\phi_n(u)$ spełnia dodatkowe ograniczenia

$$K(u,y) \leqslant K(0,y), \int_y K^2(0,y)\,dy < \infty, \int_y Q^2(y) K(0,y)\,dy < \infty, \tag{11}$$

to wyznaczony ciąg $\phi_n(u)$ charakterystyk modeli jest zbieżny do $\phi^*(u)$.

Dokładniej

$$\phi_n(u) \longrightarrow \phi^*(u), \tag{12}$$

gdy $n \to \infty$, według prawdopodobieństwa dla prawie wszystkich n [5].

Zbieżność zachodzi dla szerokiej klasy funkcji $K(u,y)$ i wielu ciągów h_n. Ich wybór należy do eksperymentatora. Przykładem ciągu h_n, który spełnia warunek (9) może być

$$h_n = cn^{\alpha},$$

przy czym

$$-\frac{1}{2k+1} < \alpha < 0, \qquad c > 0.$$

Jeśli np. $K=1$, to wejście jest skalarem a jako h_n można wybrać

$$h_n = cn^{-\frac{1}{4}}, \quad c > 0.$$

Przykłady funkcji $K(u,y)$ będą podane w dalszej części pracy.

5. P r z y k ł a d y

Jak już zaznaczono, istnieje wiele funkcji $K(u,y)$ spełniających warunki (8) i (11). Dla każdej z nich ciąg charakterystyk $\Phi_n(u)$ jest zbieżny do $\Phi^*(u)$. Umiejętny jej wybór w znacznym stopniu ułatwia rozwiązanie zadania minimalizacji (10), tzn. wyznaczenie charakterystyki modelu.

Załóżmy teraz, że:

$$Q(y) = y^2,$$

tzn., że wskaźnik jakości identyfikacji ma następującą postać:

$$\int_u \int_y \left(y - \Phi(u)\right)^2 f(u,y)\, du\, dy .$$

Niech teraz

$$K(u,y) = K_u(u)\, K_y(Y), \tag{13}$$

przy czym

$$K_y(Y) = \begin{cases} 1 & \text{jeśli } |y| \leq \frac{1}{2}, \\ 0 & \text{jeśli } |y| > \frac{1}{2}. \end{cases} \tag{14}$$

Wyrażenie (10) przyjmuje teraz następującą postać:

$$\sum_{i=1}^{n} K_u\left(\frac{u-u_i}{h_n}\right) \int_{y_i - \frac{h_n}{2}}^{y_i + \frac{h_n}{2}} \left(y - \Phi(u)\right)^2 dy$$

W celu znalezienia minimum przyrównując pochodną $d\Phi(u)$ do zera otrzymuje się po prostych przekształceniach

$$\Phi_n(u) = \frac{\sum_{i=1}^{n} y_i K_n\left(\frac{u-u_i}{h_n}\right)}{\sum_{i=1}^{n} K_n\left(\frac{u-u_i}{h_n}\right)} \tag{15}$$

Dla różnych funkcji $K_n(u)$ otrzymuje się różne warianty powyższego wzoru. Poniżej podano kilka przykładów:

1) $$K_n(u) = 2^{-K} e^{-\|u\|},$$

przy czym

$$\|u\| = \sum_{j=1}^{K} |u^{(j)}|,$$

($u^{(j)}$ oznacza j-tą współrzędną wektora u).

Wówczas

$$\Phi_n(u) = \frac{\sum_{i=1}^{n} y_i e^{-\frac{\|u-u_i\|}{h_n}}}{\sum_{i=1}^{n} e^{-\frac{\|u-u_i\|}{h_n}}}. \tag{16}$$

2) $$K_n(u) = (2\pi)^{-\frac{K}{2}} e^{-\frac{\|u\|^2}{2}},$$

przy czym

$$\|u\| = \sqrt{\sum_{j=1}^{K} |u^{(j)}|^2}$$

Zatem

$$\Phi_n(u) = \frac{\sum_{i=1}^{n} y_i e^{-\frac{\|u-u_i\|}{h_n}}}{\sum_{i=1}^{n} e^{-\frac{\|u-u_i\|}{h_n}}} \tag{17}$$

3) przy czym

$$K_n(u) = \begin{cases} 1 & \text{jeśli } \|u\| \leq \frac{1}{2}, \\ 0 & \text{jeśli } \|u\| > \frac{1}{2}, \end{cases}$$

$$\|u\| = \max |u^{(j)}|,$$

Wówczas

$$\Phi_n(u) = \frac{1}{N_n} \sum_{i_n} y_{i_n} \tag{18}$$

przy czym i_n oznacza te wskaźniki, dla których $\|u - u_i\| \leq \frac{h_n}{2}$, N_n jest ilością tych wskaźników.

Jeśli u jest skalarem, to podane wzory przyjmują znacznie prostsze formy. Dla dwóch pierwszych otrzymuje się np.

$$\phi_n(u) = \frac{\sum_{i=1}^{n} y_i \, e^{-\frac{|u-u_i|}{h_n}}}{\sum_{i=1}^{n} e^{-\frac{|u-u_i|}{h_n}}},$$

$$\phi_n(u) = \frac{\sum_{i=1}^{n} y_i \, e^{-\frac{(u-u_i)^2}{2h_n^2}}}{\sum_{i=1}^{n} e^{-\frac{(u-u_i)^2}{2h_n^2}}}.$$

Rozważmy teraz przypadek, gdy funkcja strat ma postać

$$Q(y) = |y|$$

tzn. kryterium jakości identyfikacji jest następujące:

$$\int_{u}\int_{y} |y-\phi(u)| \, f(u,y) \, du \, dy.$$

Jeśli przyjąć $K(u,y)$ według wzorów (13) i (14), to interesujące nas wyrażenie (10) sprowadzi się do następującego:

$$\sum_{i=1}^{n} K_n\left(\frac{u-u_i}{h_n}\right) \int_{y_i-\frac{h_n}{2}}^{y_i+\frac{h_n}{2}} |y-\phi(u)| \, dy.$$

Przyrównanie pochodnej $d/d\phi(u)$ do zera prowadzi do równania:

$$\sum_{i=1}^{n} K_n\left(\frac{u-u_i}{h_n}\right)\left[\left|y_i+\frac{h_n}{2}-\phi(u)\right|-\left|y_i-\frac{h_n}{2}-\phi(u)\right|\right]=0 \qquad (19)$$

Dla różnych funkcji $K_n(u)$, np. jak w przykładach (1,2 i 3) otrzymuje się różne wersje równania.

Zauważmy, że jest to równanie o jednej niewiadomej $\phi(u)$, niezależnie od wymiaru wektora u. Należy je rozwiązać w odpowiednio gęsto wybranych punktach $u \in U$. Można w tym celu stosować bardzo proste algorytmy obliczeniowe, np. reguła falsi. Pomiędzy wyznaczonymi w ten sposób wartościami charakterystyki $\phi_n(u)$ można zastosować dowolną z metod interpolacji i otrzymaną w ten sposób funkcję traktować jako wynik identyfikacji.

6. Zakończenie

W odróżnieniu od dotychczas znanych, podana w pracy metoda pozwala rozwiązać problem identyfikacji w jego pierwotnej, nieparametrycznej postaci. Dla kwadratowej funkcji strat charakterystykę $\phi_n(u)$ modelu otrzymuje się w postaci analitycznej. Dla innych, np. modułowej, w celu wyznaczenia $\phi_n(u)$ w wybranym punkcie należy rozwiązać równanie o jednej niewiadomej, co nie przedstawia istotnych trudności obliczeniowych.

7. Literatura

1. Bubnicki Z., Metody aproksymacyjne w identyfikacji obiektów, Modele matematyczne i identyfikacja procesów, Warszawa 1969
2. Bubnicki Z., System identification via estimation of probability distributions and moments, II IFAC Symposium on Identification, Prague 1970.
3. Węgrzyn S., Modele matematyczne i identyfikacja procesów, Warszawa 1969

Greblicki W., Zastosowanie metod ustalania empirycznych rozkładów prawdopodobieństwa w identyfikacji obiektów, Modele matematyczne i identyfikacja procesów, Warszawa 1969.

Greblicki W., Praca doktorska,

Parzen E., On estimation of a probability density function and mode. Ann Math. Statist. vol. 33, 1962, s. 1065-1076

Watson G. S., Leadbetter M.R., On the estimation of the probability density. Ann. Math. Statist., vol. 34, 1963, s. 480-491