Włodzimierz Greblicki
Politechnika Wrocławska

NIEPARAMETRYCZNA IDENTYFIKACJA OBIEKTÓW STATYCZNYCH

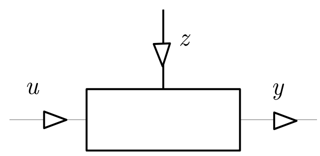
Prace V Krajowej Konferencji Automatyki,
Sekcja 1, 19–24, Gdańsk 1971

1. Wstęp

Znane metody rozwiązania probabilistycznego problemu identyfikacji stacjonarnego obiektu statycznego polegają na przyjęciu klasy charakterystyk modeli, w której następnie minimalizuje się wskaźnik jakości identyfikacji. Niekiedy nieznany, łączny rozkład prawdopodobieństwa wielkości wejściowych i wyjściowych aproksymuje się funkcjami z wybranej klasy i otrzymane przybliżenie traktuje się 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 identyfikacji pozwala uniknąć tych błędów. Wyznacza ona model najlepszy spośród wszystkich modeli.

2. Założenia

Przez u będzie oznaczone wejście, a przez y wyjście badanego obiektu o schemacie blokowym jak na rys. 1. Wektor wejściowy u należy do k-wymiarowej przestrzeni $\mathcal{U}$, a skalar y do przestrzeni $\mathcal{Y}$ 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



Rysunek 1: Schenat blokowy badanego obiektu.

$$(u_1, y_1), (u_2, y_2), \dots, (u_n, y_n) \tag{1}$$

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).

3. Problem identyfikacji

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

$$\bar{Q}(\varphi(u)) = \int\int Q(y-\varphi(u))^2 f(u,y)dudy, \tag{2}$$

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

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

$$\int\int Q(y-\bar{\varphi}(u;a))^2 \bar{f}(u,y;b)dudy. \tag{3}$$

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 $\varphi_1(u;b^*)$ wyznacza się poprzez minimalizację względem φ wyrażenia

$$\int\int Q(y-\varphi(u))^2 \bar{f}(u,y;b^*)dudy, \tag{5}$$

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 $\bar{\varphi}(u;a)$ i aproksymacja nieznanego rozkładu, [2], [4]. Dla wyznaczenia modelu należy rozwiązać zadanie minimalizacji

$$\min_a \int\int Q(y-\bar{\varphi}(u;a)^2 \bar{f}(u,y;b^*)dudy. \tag{6}$$

Oznaczmy przez $\bar{\varphi}(u;a^*,b^*)$ otrzymaną w ten sposób charakterystykę modelu.

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

I. Problem nieparametryczny

II. Klasa modeli określona parametrycznie

III. Parametryczna aproksymacja gęstości

IV. Klasa modeli określona parametrycznie
Parametryczna aproksymacja gęstości

Z oczywistych nierówności

$$\bar{Q}(\varphi^*(u)) \le \bar{Q}(\bar{\varphi}(u;a^*)) \le \bar{Q}(\bar{\varphi}(u;a^*,b^*)),$$

$$\bar{Q}(\varphi^*(u)) \le \bar{Q}(\varphi_1(u;b^*)) \le \bar{Q}(\bar{\varphi}(u;a^*,b^*))$$

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

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

Sposoby rozawiązania zagadnień II, III i IV są ogólnie znane [1]–[3]. W tej pracy podano metodę rozwiązania neparametrycznego problemu I. Pozwala ona wyznaczyć ciąg charakterystyk $\varphi_n(u)$ zbieżnych do $\varphi^*(u)$.

4. Rozwiązanie problemu nieparametrycznego

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

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

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

$$0 \le K(u,y) < \infty, \int\int K(u,y)dudy = 1, \int\int K^2(u,y)dudy < \infty \tag{8}$$

i

$$h_n > 0, h_n \to 0, nh_n \to 0, \tag{9}$$

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

$$\lim_{n\to\infty} E \int\int (f_n(u,y) - f(u,y))^2 dudy = 0.$$

Podany ciąg gęstości empirycznych jest więc zbieżny do gęstości rzeczywistej.

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

$$\int\int Q(y-\varphi(u))f_n(u,y)dudy,$$

lub, co na jedno wychodzi,

$$\int Q(y-\varphi(u))f_n(u,y)dy. \tag{10}$$

Po uwzględnieniu wzoru (7), powyższe wyrażenie przyjmuje postać

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

Jeśli teraz funkcja $\varphi_n(u)$ spełnia dodatkowe ograniczenia

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

to wyznaczony ciąg $\varphi_n(u)$ charakterystyk modeli jest zbieżny do $\varphi^*(u)$. Dokładniej,

$$\varphi_n(u)\to\varphi^*(u), \tag{12}$$

gdy $n\to\infty$ według prawdopodonieństwa, [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

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

Jeśli np. $k=1$, to wejście jest skalarem a jako h_n nożna wybrać

$$h_n = cn^{-1/4}, c>0.$$

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

5. Przykłady

Jak już zaznaczono, istnieje wiele funkcji $K(u,y)$ spełniających warunki (8) i (11). Dla każdej z nich ciąg charakterystyk $\phi_n(u)$ zbieżnych 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 wskaźnikjakości identyfikacji ma następującą postać:

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

Niech teraz

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

przy czym

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

Wyrażenie (10) przyjmuje teraz następującą postać:

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

W celu znalezienia minimum, przyrównując pochodną $d/d\phi(u)$ do zera, otrzymuje się po prostych przekształceniach:

$$\phi_n(u) = \frac{\displaystyle\sum_{i=1}^{n} y_i K_u \left(\frac{u - u_i}{h_n} \right)}{\displaystyle\sum_{i=1}^{n} K_u \left(\frac{u - U_i}{h_n} \right)}.$$

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

a)

$$K_u(u) = \frac{1}{2^k} e^{-||u||},$$

przy czym $||u|| = \sum_{j=1}^{k} |u^{(j)}|$ ($u^{(j)}$ oznacza j-tą współrzędną wektora u). Wówczas

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

b)

$$K_u(u) = \frac{1}{(2\pi)^{k/2}} e^{-||u||^2/2},$$

przy czym $||u|| = \sqrt{\sum_{j=1}^{k} |u^{(j)}|^2}$. Zatem

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

c)

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

przy czym $||u|| = \max_j |u^{(j)}|$. Wówczas

$$\phi_n(u) = \frac{1}{N_n} \sum_{j=1}^{n} y_{in},$$

przy czym $_{in}$ oznacza te wskaźniki, dla których $||u - U_i|| \leq h_n/2$, N_n jest ilością tych wskaźników.

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

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

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

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:

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

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

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

Dla różnych funkcji $K_u(u)$ (np. jak w przykładach a), b) i c)) 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. regułę 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 innnych, np. modułowej, w celu wyznaczenia $\phi_n(u)$ w wybranym punkcie należy rozwiązać równanie o jednej niewiadome, co nie przedstawia istotnych trudności obliczeniowych.

Literatura

[1] Bubnicki Z., "Metody aproksymacyjne w identyfikacji obiektów", [w] *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 obiektów, Warszawa, 1969.

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

[5] Greblicki W., "Zastosowanie metod określania empirycznych rozkładów prawdopodobieństwa w identyfikacji obiektów", (rozprawa doktorska), Instytut Cybernetyki Technicznej, Politechnika Wrocławska, Wrocław 1970.

[6] Parzen E., "On estimation of a probability density function and mode", *Ann. Math. Statist.*, vol. 17, 1065–1076, 1962.

[7] Watson G. S., Leadbetter M. R., "On the estimation of the probability density, I", *Ann. Math. Statist.*, vol. 34, 480–491, 1963.