

# 4

## Estymatory jądrowe w zagadnieniach badań systemowych

---

*Piotr Kulczycki*<sup>1</sup>

### **Streszczenie**

W rozdziale tym zaprezentowano koncepcję statystycznych estymatorów jądrowych. Dzięki możliwościom i powszechności współczesnej techniki komputerowej ten rodzaj estymatorów staje się obecnie podstawową metodą estymacji nieparametrycznej, umożliwiającą wyznaczanie charakterystyk zmiennych losowych bez arbitralnych założeń dotyczących przynależności ich rozkładu do ustalonej klasy. Najpierw przedstawiono podstawy metodyki konstruowania estymatorów jądrowych, a następnie ich możliwości aplikacyjne zilustrowano na przykładzie zagadnień współczesnych badań systemowych. Przedstawiono cztery problemy, z zakresu archeologii, medycyny, sterowania automatycznego i telekomunikacji, rozważając kolejno zadania rozpoznania elementów nietypowych, wyostrzania informacji nieprecyzyjnej, identyfikacji parametrycznej, a także określenia przestrzennego rozkładu popytu.

**Słowa kluczowe:** analiza danych, estymacja nieparametryczna, estymatory jądrowe, badania systemowe, element nietypowy (odosobniony), informacja nieprecyzyjna, identyfikacja modelu, rozkład popytu.

---

<sup>1</sup> Instytut Badań Systemowych PAN, ul. Newelska 6, 01-447 Warszawa;  
e-mail: kulczycki@ibspan.waw.pl

Politechnika Krakowska, Zakład Automatyki (E-32), ul. Warszawska 24, 31-155 Kraków;  
e-mail: kulczycki@pk.edu.pl

## 4.1. Wstęp

Jednym z podstawowych problemów współczesnej nauki i złożonych działań praktycznych jest wyznaczenie funkcji charakteryzujących badany obiekt, na podstawie dostępnych danych eksperymentalnych. I tak na przykład funkcję regresji, opisującą zależność między dwoma skorelowanymi wielkościami, określa się, wykorzystując otrzymane doświadczalnie pary tych wielkości. Innym typowym zagadnieniem jest, estymacja gęstości rozkładu probabilistycznego zmiennej losowej na podstawie uzyskanej z niej próby.

Klasyczne procedury polegają na arbitralnym założeniu postaci badanych funkcji, a potem wyznaczeniu definiujących je parametrów. Określa się je mianem **metod parametrycznych**. Ich istotnymi zaletami są prostota teoretyczna oraz obliczeniowa, a także powszechna znajomość i dostępność w literaturze. Obecnie, w miarę gwałtownego rozwoju techniki komputerowej, coraz częściej są wprowadzane **metody nieparametryczne**, których zasadniczą cechą jest brak arbitralnego założenia o postaci badanej funkcji. W ujęciu probabilistycznym podstawowym przedstawicielem tych metod stają się estymatory jądrowe. Chociaż ich koncepcja jest prosta, a interpretacja przejrzysta, to praktyczne zastosowania nie są możliwe bez użycia wysokiej klasy komputerów, co jeszcze w niedalekiej przeszłości znacząco utrudniało badania teoretyczne, a zwłaszcza aplikacyjne.

W niniejszym rozdziale zostały przedstawione podstawy metodyki estymatorów jądrowych, a ich praktyczne możliwości są zilustrowane na przykładzie zagadnień współczesnych badań systemowych. Ta ogólna i interdyscyplinarna dziedzina wydaje się szczególnie predestynowana do zastosowań estymatorów jądrowych z uwagi na ich uniwersalne możliwości.

Rozdział ten stanowi zbiór podstawowych informacji zawartych w **monografii** [18], która ukazała się niedawno na rynku księgarskim. Estymatory jądrowe nie doczekały się uprzednio opracowania w języku polskim, oprócz krótkich wzmianek w nielicznych podręcznikach, na przykład [4, 6], lub wręcz nawiązań do samej ich koncepcji [7, 9]. I tak, podrozdział 4.2 zawiera przegląd metodyki konstruowania tego typu estymatorów, z ukierunkowaniem na aspekty aplikacyjne, które w podrozdziale 4.3 zostaną wykorzystane do wybranych zagadnień badań systemowych.

## 4.2. Metodyka konstruowania estymatorów jądrowych

Niech dana będzie  $n$ -wymiarowa zmienna losowa  $X$ , której rozkład ma gęstość  $f$ . Jej **estymator jądrowy**  $\hat{f} : \mathbb{R}^n \rightarrow [0, \infty)$  wyznacza się na podstawie wartości  $m$ -elementowej próby losowej

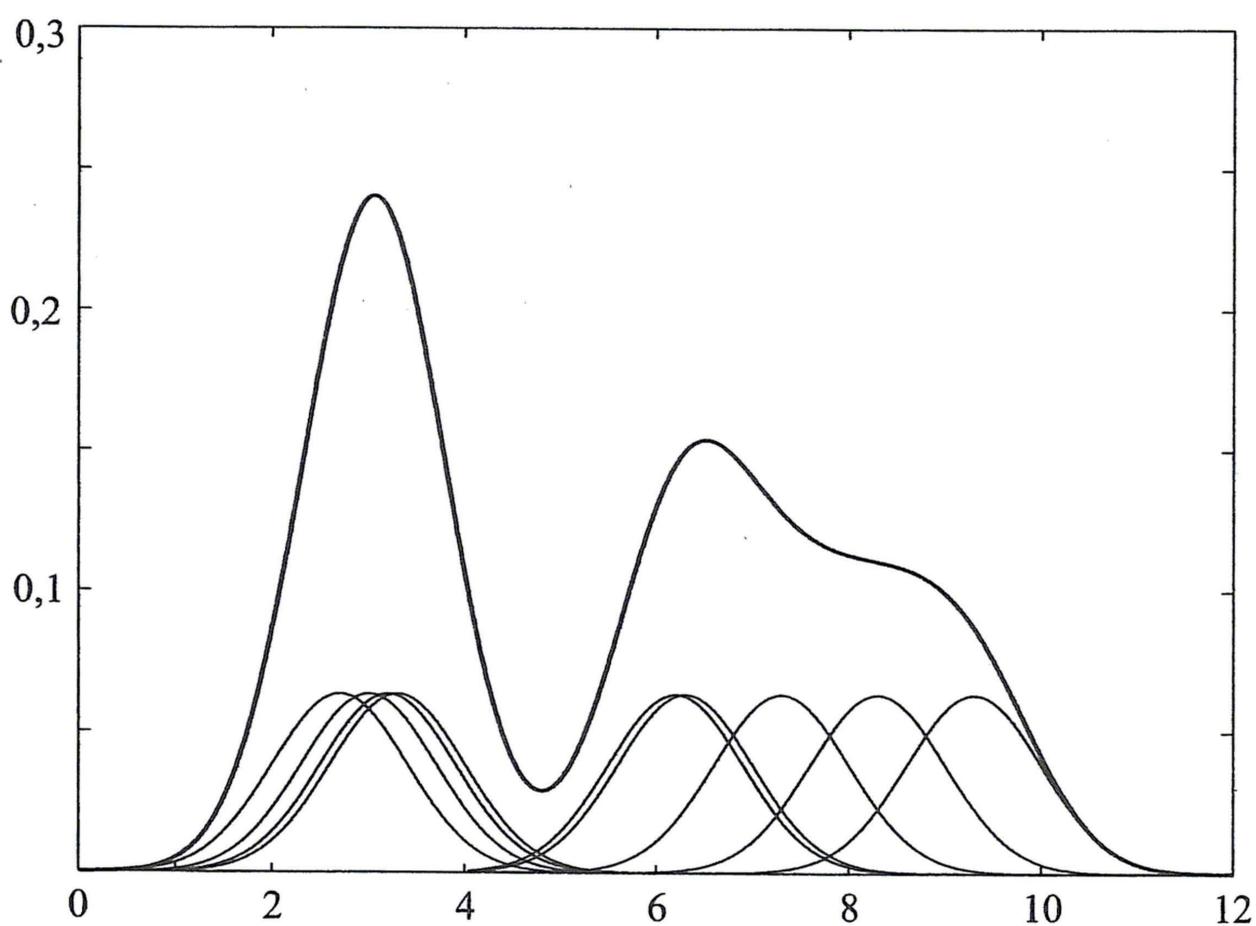
$$x_1, x_2, \dots, x_m, \tag{4.1}$$

uzyskanej ze zmiennej  $X$ , i w swej podstawowej postaci jest on zdefiniowany wzorem

$$\hat{f}(x) = \frac{1}{mh^n} \sum_{i=1}^m K\left(\frac{x - x_i}{h}\right), \quad (4.2)$$

gdzie mierzalna, symetryczna względem zera oraz mająca w tym punkcie słabe maksimum globalne funkcja  $K : \mathbb{R}^n \rightarrow [0, \infty)$  spełnia warunek  $\int_{\mathbb{R}^n} K(x)dx = 1$  i jest nazywana **jądrem**, natomiast dodatni współczynnik  $h$  określa się mianem **parametru wygładzania**.

Interpretację powyższej definicji ilustruje rysunek 4.1, na przykładzie zmiennej losowej jednowymiarowej, czyli gdy  $n = 1$ , dla 9-elementowej próby losowej, a zatem  $m = 9$ . W przypadku pojedynczej realizacji  $x_i$ , funkcja  $K$  (przesunięta o wektor  $x_i$  oraz przeskalowana współczynnikiem  $h$ ) reprezentuje oszacowanie rozkładu zmiennej losowej  $X$  po otrzymaniu wartości  $x_i$ . Dla  $m$  niezależnych realizacji  $x_1, x_2, \dots, x_m$  oszacowanie to przyjmuje postać sumy takich pojedynczych oszacowań. Współczynnik  $1/mh^n$  normuje uzyskaną funkcję w celu zagwarantowania warunku  $\int_{\mathbb{R}^n} \hat{f}(x)dx = 1$ , jaki powinna spełniać gęstość rozkładu probabilistycznego.



Rys. 4.1. Jądrowy estymator gęstości (4.2) rozkładu jednowymiarowej zmiennej losowej

Warto zauważyć, że estymator jądrowy umożliwia oszacowanie gęstości praktycznie **dowolnego rozkładu**, bez żadnych założeń o jego przynależności do ustalonej klasy. Nietypowe, złożone rozkłady (w tym także wielomodalne) są tu traktowane tak samo jak podręcznikowe rozkłady jednomodalne. W przypadku wielowymiarowym, czyli przy  $n > 1$ , pozwala to między innymi na identyfikację wszechstronnych zależności między poszczególnymi współrzędnymi badanej zmiennej losowej.

Ustalenie wielkości wprowadzonych w definicji (4.2), a zatem wybór postaci jądra  $K$  oraz wyznaczenie wartości parametru wygładzania  $h$ , jest najczęściej dokonywane na podstawie kryterium minimum scałkowanego błędu średniokwadratowego. Zagadnienia te będą przedmiotem rozważań punktów 4.2.1 i 4.2.2.

### 4.2.1. Wybór postaci jądra

Ze statystycznego punktu widzenia **postać jądra** okazuje się nie mieć istotnego znaczenia i dzięki temu pojawia się możliwość, aby wybór funkcji  $K$  był arbitralny, uwzględniający przede wszystkim pożądane własności otrzymanego estymatora, na przykład klasę jego regularności, przyjmowanie dodatnich wartości lub też inne cechy istotne dla konkretnego problemu, w tym zwłaszcza dogodność obliczeniową. Ma to szczególne znaczenie w przypadku zagadnień badań systemowych, gdzie estymator jądrowy stanowi najczęściej podstawę do dalszych wszechstronnych i złożonych rozważań przedmiotowych, a uzyskane własności otrzymanego estymatora jądrowego mogą nie tylko ułatwiać dalszą procedurę, lecz często wręcz umożliwić uzyskanie konkretnego, dogodnego do zastosowań wyniku.

W przypadku jednowymiarowym jako funkcję  $K$  przyjmuje się klasyczne postacie gęstości rozkładów probabilistycznych, na przykład gęstość rozkładu normalnego, Cauchy'ego, trójkątnego i inne, lub też – w szczególnych zagadnieniach – ich kombinacje liniowe. Najefektywniejszym w sensie kryterium błędu średniokwadratowego jest tak zwane **jądro Epanecznikowa**

$$K(x) = \begin{cases} \frac{3}{4}(1 - x^2) & \text{dla } x \in [-1, 1], \\ 0 & \text{dla } x \in (-\infty, -1) \cup (1, \infty). \end{cases} \quad (4.3)$$

W przypadku wielowymiarowym stosuje się dwa naturalne uogólnienia powyższej koncepcji, a mianowicie **jądro radialne**

$$K(x) = CK \left( \sqrt{x^T x} \right) \quad (4.4)$$

oraz **jądro produktowe**

$$K(x) = K([x_1, x_2, \dots, x_n]^T) = \mathcal{K}(x_1) \cdot \mathcal{K}(x_2) \cdot \dots \cdot \mathcal{K}(x_n), \quad (4.5)$$

gdzie  $\mathcal{K}$  oznacza omówione uprzednio jądro jednowymiarowe, a  $C$  jest dodatnią stałą wyznaczoną tak, aby spełniony był warunek  $\int_{\mathbb{R}^n} K(x) dx = 1$ . Dla dowolnie ustalonego jądra jednowymiarowego  $\mathcal{K}$  bardziej efektywne jest jądro radialne (4.4) niż produktowe (4.5), lecz różnica – z aplikacyjnego punktu widzenia – jest nieznaczna. Fakt ten powoduje, że w praktycznych zastosowaniach często preferuje się jądro produktowe. Poza szczególnymi zastosowaniami statystycznymi jest ono bowiem znacznie dogodniejsze w dalszej analizie – na przykład procedury całkowania i różniczkowania jądra produktowego niewiele się różnią od przypadku jednowymiarowego. Wśród jąder radialnych najefektywniejsze jest radialne jądro Epanecznikowa, czyli definiowane zależnością (4.4), gdy  $\mathcal{K}$  jest dane

wzorem (4.3). Także w rodzinie jąder produktowych najefektywniejsze okazuje się produktowe jądro Epanecznikowa, określone przez równości (4.5) oraz (4.3).

**Podsumowując:** możliwość znacznej elastyczności przy doborze postaci jądra  $K$  stanowi istotną praktyczną zaletę, ujawniającą się proporcjonalnie do złożoności konkretnego zastosowania.

#### 4.2.2. Określenie wartości parametru wygładzania

W przeciwieństwie do postaci jądra przyjęta **wartość parametru wygładzania** ma istotny wpływ na jakość otrzymanego estymatora jądrowego. Opracowane jednak zostały dogodne algorytmy umożliwiające automatyczne obliczenie zbliżonej do optymalnej – w sensie średniokwadratowym – wartości  $h$ , na podstawie próby losowej (4.1).

I tak, jako przybliżoną można przyjąć, w początkowej fazie analizy, wartość wyznaczoną dla rozkładu normalnego:

$$h = \left( \frac{V(K)}{U(K)^2} \pi^{n/2} \frac{2^{n+2}}{n+2} \frac{1}{m} \right)^{1/(n+4)} \bar{\sigma}, \quad (4.6)$$

gdzie  $\bar{\sigma}$  oznacza średnią geometryczną odchyłeń standardowych poszczególnych współrzędnych zmiennej  $X$ , natomiast  $V(K) = \int_{\mathbb{R}^n} K(y)^2 dy$  oraz  $U(K) = \int_{\mathbb{R}^n} y^T y K(y) dy$ . Uniwersalna jest **metoda krzyżowego uwiarygodnienia** (*cross-validation*), w ramach której wyznacza się wartość realizującą minimum rzeczywistej funkcji  $g : (0, \infty) \rightarrow \mathbb{R}$  zdefiniowanej równością

$$g(h) = \frac{1}{m^2 h^n} \sum_{i=1}^m \sum_{j=1}^m \tilde{K} \left( \frac{x_j - x_i}{h} \right) + \frac{2}{m h^n} K(0), \quad (4.7)$$

przy czym  $\tilde{K}(x) = K^{*2}(x) - 2K(x)$ , gdzie  $K^{*2}$  oznacza kwadrat splotowy funkcji  $K$ . Stosując w celu poszukiwania minimum metody iteracyjne, wartość początkową procedury można wyznaczyć ze wzoru (4.6).

Dla szczególnych przypadków opracowano także szereg dogodniejszych algorytmów, przykładowo prostą i efektywną **metodę podstawień** (*plug-in*) stosowaną do przypadku jednowymiarowego.

#### 4.2.3. Dodatkowe procedury

Zaprezentowana powyżej podstawowa postać estymatora jądrowego (4.2) może być uogólniona w celu generalnej poprawy jego własności oraz ewentualnie uzupełniana o dodatkowe aspekty dopasowujące model do badanej rzeczywistości. Jako przykłady **uogólnień** zostaną kolejno przedstawione procedury

- modyfikacji parametru wygładzania (punkt 4.2.4),
- transformacji liniowej (punkt 4.2.5),

natomiast w ramach **uzupełnień** zaprezentowane będą

- ograniczenie nośnika (punkt 4.2.6),
- uwzględnienie współrzędnych binarnych (punkt 4.2.7).

#### 4.2.4. Modyfikacja parametru wygładzania

Wartość parametru wygładzania  $h$ , wprowadzonego w definicji (4.2), jest taka sama dla wszystkich jąder przyporządkowanych poszczególnym wartościom próby losowej (4.1). Ogólnie, mała wartość tego parametru powoduje „wyszczuplenie”, natomiast duża – „spłaszczenie” jądra. Jeżeli zatem możliwe jest zindywidualizowanie wpływu parametru wygładzania na poszczególne jądra, to w rejonach „zagęszczenia” elementów próby losowej wartość ta powinna być zmniejszona (co pozwoli lepiej ukazać specyficzne cechy rozkładu), w przeciwieństwie do obszarów, w których elementy te są „rzadkie”, gdzie powinna ulec zwiększeniu (co spowoduje dodatkowe wygładzenie tak zwanych „ogonów”). Służąca temu celowi **modyfikacja parametru wygładzania** realizuje powyższy cel zgodnie z następującym algorytmem.

Najpierw wyznacza się podstawową postać estymatora jądrowego (4.2), po czym obliczane są jego wartości dla poszczególnych elementów próby losowej, czyli  $\hat{f}(x_1), \hat{f}(x_2), \dots, \hat{f}(x_m)$ , a także ich średnia geometryczna  $\bar{s}$ . **Parametry modyfikujące**  $s_i > 0$  ( $i = 1, 2, \dots, m$ ) określa się wzorem

$$s_i = \left( \frac{\hat{f}(x_i)}{\bar{s}} \right)^{-c}, \quad (4.8)$$

przy czym  $c \geq 0$ , i w konsekwencji estymator jądrowy ze zmodyfikowanym parametrem wygładzania dany jest jako

$$\hat{f}(x) = \frac{1}{mh^n} \sum_{i=1}^m \frac{1}{s_i^n} K \left( \frac{x - x_i}{hs_i} \right). \quad (4.9)$$

Powyższy wzór jest uogólnieniem definicji (4.2). Jeżeli bowiem przyjąć  $c = 0$ , to  $s_i \equiv 1$  i zależności te są równoważne. Parametr  $c$  stanowi o intensywności procedury modyfikacji. Wnioski wynikające z kryterium minimum błędu średniokwadratowego wskazują na wartość  $c = 0,5$ .

Estymatory jądrowe z modyfikacją parametru wygładzania mają w praktycznych zastosowaniach – poza oczywistym polepszeniem jakości oszacowania – wiele dodatkowych **pozytywnych własności**. Przede wszystkim są one odporniejsze na nieprecyzyjny dobór wartości parametru wygładzania  $h$ . Co więcej, stosowanie tej procedury zmniejsza różnicę efektywności poszczególnych typów jąder względem optymalnego jądra Epanecznikowa (4.3), podobnie jak w przypadku wielowymiarowym zmniejsza różnicę efektywności jądra produktowego (4.5) względem radialnego (4.4). Własności te są szczególnie cenne w praktycznych zastosowaniach, gdyż dodatkowo zwiększają możliwość preferencji jąder korzystnych dla dalszej analizy w konkretnych zadaniach aplikacyjnych.

#### 4.2.5. Transformacja liniowa

W przypadku wielowymiarowym, gdy jest używane jądro radialne, można zastosować prostą procedurę dopasowującą kształt jądra do postaci rozkładu tej zmiennej, opartą na koncepcji **transformacji liniowej**. Definicję jądra radialnego

(4.4) uogólnia się wtedy do wzoru

$$K(x) = \frac{C}{\sqrt{\det(R)}} \mathcal{K} \left( \sqrt{x^T R^{-1} x} \right), \quad (4.10)$$

przy czym  $R$  jest macierzą dodatnio określoną, natomiast znaczenie i wartość stałej  $C$  pozostają niezmienione. Jeżeli  $R$  jest macierzą jednostkową, to powyższa postać jest równoważna zależności (4.4). Szczególnie korzystne rezultaty otrzymuje się, przyjmując

$$R = C\hat{ov}, \quad (4.11)$$

gdzie  $C\hat{ov}$  oznacza estymator macierzy kowariancji rozkładu zmiennej  $X$ .

#### 4.2.6. Ograniczenie nośnika

W praktycznych zagadnieniach poszczególne współrzędne zmiennej losowej  $X$  mogą reprezentować różnorodne wielkości, zwłaszcza w przypadku tak interdyscyplinarnej dyscypliny jak badania systemowe. Wiele z nich, między innymi wielkości związane z długością lub czasem, jeśli mają być poprawnie i ściśle interpretowane, powinny należeć jedynie do ograniczonych podzbiorów, na przykład zbioru liczb nieujemnych. Aby uniknąć wynikłych stąd błędów obliczeniowych i nieporozumień interpretacyjnych, można zastosować prostą **procedurę ograniczającą nośnik** estymatora jądrowego.

Najpierw przedstawiony zostanie przypadek jednowymiarowej zmiennej losowej i jej **ograniczenia lewostronnego**, czyli gdy wymagany jest warunek  $\hat{f}(x) = 0$  dla  $x < x_*$  przy ustalonym  $x_* \in \mathbb{R}$ . Fragment dowolnego  $i$ -tego jądra, który znajduje się poza przedziałem  $[x_*, \infty)$ , zostaje tu „odbity” symetrycznie względem ograniczenia  $x_*$  i jest traktowany jako fragment jądra „zaczepionego” w symetrycznym „odbiciu” elementu  $x_i$  względem ograniczenia  $x_*$ , czyli w punkcie  $2x_* - x_i$ . Jeśli zatem zdefiniować funkcję  $K_{x_*} : \mathbb{R} \rightarrow [0, \infty)$  wzorem

$$K_{x_*}(x) = \begin{cases} K(x), & \text{gdy } x \geq x_*, \\ 0, & \text{gdy } x < x_*, \end{cases} \quad (4.12)$$

to podstawową postać estymatora jądrowego (4.2) można wtedy określić następująco:

$$\hat{f}(x) = \frac{1}{mh^n} \sum_{i=1}^m \left[ K_{x_*} \left( \frac{x - x_i}{h} \right) + K_{x_*} \left( \frac{x + x_i - 2x_*}{h} \right) \right]. \quad (4.13)$$

Fragmenty poszczególnych jąder „ucinane” poza założonym nośnikiem są zatem „uzupełniane” wewnątrz nośnika w bezpośrednim sąsiedztwie ograniczenia, a więc w zakresie akceptowalnego najczęściej w praktyce błędu.

Analogicznie rozważa się przypadek **prawostronnego ograniczenia** nośnika estymatora jądrowego. W przypadku wielowymiarowym powyższą koncepcję można stosować odrębnie wobec poszczególnych współrzędnych badanej zmiennej losowej.

### 4.2.7. Współrzędne binarne

W wielu aplikacyjnych zagadnieniach badań systemowych występują **wielkości binarne**, to znaczy przyjmujące jedynie dwie wartości, symbolicznie oznaczane dalej jako 0 oraz 1. Na przykład istotną informacją w bazie danych medycznych jest płeć pacjenta dwoistej postaci „kobieta lub mężczyzna”. Poza wielkościami mającymi ze swej natury postać binarną często stosuje się także uproszczenia do takiej formy opisu nawet złożonych zjawisk, lecz o pomniejszym wpływie na ostateczny wynik, wprowadzając wobec nich postać dwuwartościową, na przykład typu „wpływ pozytywny lub negatywny”. Metodyka estymatorów jądrowych umożliwia uwzględnienie wielkości binarnych.

Rozważana najpierw będzie  $k$ -wymiarowa binarna zmienna losowa  $Y : \Omega \rightarrow \{0, 1\}^k$ . Jej rozkład jest charakteryzowany przez  $2^k$  prawdopodobieństw wystąpienia każdego z możliwych  $k$ -wymiarowych wektorów o wartościach binarnych. W wielu zagadnieniach aplikacyjnych o wartości prawdopodobieństwa wystąpienia danego wektora można również wnioskować, posiłkując się obserwacjami wektorów do niego „podobnych”. Im to „podobieństwo” jest większe, tym wnioskowanie takie staje się bardziej wiarygodne. Niech zatem dana będzie funkcja  $p : \{0, 1\}^k \rightarrow [0, 1]$ , przyporządkowująca każdemu  $k$ -wymiarowemu wektorowi o wartościach binarnych prawdopodobieństwo jego wystąpienia. Estymator jądrowy tej funkcji  $\hat{p} : \{0, 1\}^k \rightarrow [0, 1]$  wyznacza się na podstawie wartości  $m$ -elementowej próby losowej

$$y_1, y_2, \dots, y_m \quad (4.14)$$

uzyskanej ze zmiennej  $Y$  i jest on definiowany jako

$$\hat{p}(y) = \frac{1}{m} \sum_{i=1}^m L(y, y_i), \quad (4.15)$$

gdzie funkcję  $L : \{0, 1\}^k \times \{0, 1\}^k \rightarrow [0, 1]$  określa zależność

$$L(y, y_i) = \delta^{k-d(y, y_i)} (1 - \delta)^{d(y, y_i)}, \quad (4.16)$$

przy czym  $0,5 \leq \delta \leq 1$  oraz przyjmuje się  $0^0 = 1$ , a funkcja  $d : \{0, 1\}^k \times \{0, 1\}^k \rightarrow \mathbb{N}$  jest dana wzorem  $d(y, y_i) = (y - y_i)^T (y - y_i)$ . Wartość  $k - d(y, y_i)$  wskazuje liczbę współrzędnych, na których wektory  $y$  oraz  $y_i$  są równe, i reprezentuje wspomniane wcześniej „podobieństwo” wektorów binarnych. Funkcja  $L$  odgrywa rolę, jaką w definicji (4.2) odgrywało jądro  $K$ , i jest nazywana **jądrem binarnym**, a parametr  $\delta$  – **binarnym parametrem wygładzania**. W praktyce jego wartość wyznacza się, minimalizując funkcję  $g : [0,5; 1] \rightarrow \mathbb{R}$  definiowaną wzorem

$$g(\delta) = - \sum_{i=1}^m \log[\hat{p}_{-i}(y_i)], \quad (4.17)$$

gdzie  $\hat{p}_{-i}$  oznacza estymator (4.15) wyznaczony na podstawie próby losowej  $y_1, y_2, \dots, y_{i-1}, y_{i+1}, \dots, y_m$ , a zatem z pominięciem  $i$ -tego elementu próby (4.14).

I wreszcie, zostanie teraz rozpatrzona  $(n + k)$ -wymiarowa zmienna losowa  $Z \equiv [X, Y]^T$  będąca zestawieniem rozważanej dotychczas  $n$ -wymiarowej ciągłej



zmiennej  $X$  oraz określonej powyżej  $k$ -wymiarowej zmiennej binarnej  $Y$ . Jeżeli estymator jądrowy  $\hat{f} : \mathbb{R}^n \times \{0, 1\}^k \rightarrow [0, \infty)$  zmiennej  $Z$  jest wyznaczany na podstawie  $m$ -elementowej próby losowej

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}, \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}, \dots, \begin{bmatrix} x_m \\ y_m \end{bmatrix}, \quad (4.18)$$

to podstawową postać estymatora jądrowego (4.2) definiuje się wtedy następująco:

$$\hat{f}(x, y) = \frac{1}{mh^n} \sum_{i=1}^m K\left(\frac{x - x_i}{h}\right) L(y, y_i). \quad (4.19)$$

### 4.2.8. Liczność próby

Ostatnim parametrem wymagającym komentarza jest niezbędna **liczność próby**, a szczególnie jej zależność od wymiaru badanej zmiennej losowej. Liczność próby  $m_*$  konieczną do zapewnienia 10-procentowej dokładności w punkcie zero dla  $n$ -wymiarowego standardowego rozkładu normalnego można w przybliżeniu przyjąć jako  $m_* = 4^n$ . Ze względu na szczególną regularność powyższego rozkładu oraz znaczny liberalizm przyjętego kryterium wartości te wydają się stanowić bezwzględne minimum (sugeruje to np.  $m_* = 4$  dla  $n = 1$ ). W przypadku obecności współrzędnych binarnych powyższą wartość należy pomnożyć przez  $(3/2)^k$ . Z kolei uzyskany wynik mnoży się przez heurystycznie określone współczynniki związane z koniecznością polepszenia jakości estymacji, wielomodalnością i niesymetrią rozkładu oraz korelacją elementów próby losowej – w praktyce iloczyn tych współczynników jest równy najczęściej 3–10, a w ekstremalnych przypadkach nawet 100. Dla jednowymiarowej zmiennej losowej wymagana liczność próby wynosi w praktyce 20–50, odpowiednio zwiększając się wraz ze wzrostem wymiaru zmiennej. Jednak dzięki współczesnej technice komputerowej, nawet w złożonych zagadnieniach wielowymiarowych i przy niesprzyjających cechach rozkładów, nie musi to obecnie stanowić istotnej przeszkody aplikacyjnej. Zawsze należy przecież brać pod uwagę znaczące korzyści wynikające ze stosowania estymatorów jądrowych. Umożliwiają one bowiem identyfikację praktycznie dowolnego występującego w zadaniach aplikacyjnych rozkładu, chociaż wymagają stosownej liczności próby, adekwatnej do mnogości i wszechstronności zawartej w niej informacji.

### 4.2.9. Uwagi i komentarze

W podrozdziale 4.2 przedstawiono zarys metodyki konstruowania estymatorów jądrowych. Sformułowana została definicja podstawowej postaci tego typu estymatora (4.2) i procedury wyznaczania występujących tam wielkości. Postać ta została uogólniona przez wprowadzenie koncepcji modyfikacji parametru wygładzania (4.9) i transformacji liniowej (4.10), polepszających własności otrzymanego estymatora. Zaproponowano także dodatkowe procedury umożliwiające ograniczenie jego nośnika (14.3) oraz uwzględnienie wielkości binarnych (4.19), co niejednokrotnie pozwala lepiej dopasować uzyskany model do badanej rzeczywistości.

ści. Powyższe uogólnienia i uzupełnienia zostały sformułowane w odniesieniu do podstawowej postaci estymatora jądrowego, co dopuszcza dowolne ich łączenie, zgodnie z uwarunkowaniami konkretnego problemu aplikacyjnego. Ze względu na wyrazistą interpretację możliwe jest także wprowadzanie dalszych uogólnień. Przykładowo – na potrzeby dalszych rozważań przedstawionych w podrozdziale 4.3 – podstawową postać (4.2) można uzupełnić o nieujemne, nie wszystkie równe zero współczynniki  $w_i$  przy  $i = 1, 2, \dots, m$ , które interpretuje się jako „znaczenie” poszczególnych elementów próby losowej (4.1). Wtedy

$$\hat{f}(x) = \frac{1}{h^n \sum_{i=1}^m w_i} \sum_{i=1}^m w_i K \left( \frac{x - x_i}{h} \right). \quad (4.20)$$

Jeżeli  $w_i \equiv 1$ , to powyższa postać jest równoważna definicji (4.2).

Estymatory jądrowe umożliwiają oszacowanie gęstości rozkładu – podstawowej charakterystyki funkcyjnej jedno- i wielowymiarowej zmiennej losowej – z praktycznego punktu widzenia niezależnie od jej postaci i cech. W konsekwencji stanowi to podstawę do wyznaczenia innych charakterystyk funkcyjnych oraz parametrów. Na przykład, jeśli w przypadku jednowymiarowym zostanie wybrane takie jądro  $K$ , aby istniała analityczna postać jego funkcji pierwotnej  $I(x) = \int_{-\infty}^x K(y)dy$ , to można zdefiniować **jądrowy estymator dystrybuanty**

$$\hat{F}(x) = \frac{1}{m} \sum_{i=1}^m I \left( \frac{x - x_i}{h} \right). \quad (4.21)$$

Z kolei, jeżeli jądro  $K$  ma dodatnie wartości, to rozwiązanie równania

$$\hat{F}(x) = r \quad (4.22)$$

stanowi **jądrowy estymator kwantyla**  $\hat{q}$  stopnia  $r \in (0, 1)$ .

Analogicznie do ogólnej idei estymatora gęstości rozkładu zmiennej losowej formułowana jest koncepcja **jądrowego estymatora gęstości widmowej**, a także – szczególnie interesującego z aplikacyjnego punktu widzenia – **jądrowego estymatora funkcji regresji**. Zgodnie z generalną zasadą dotyczącą estymatorów jądrowych funkcja ta jest wyznaczana bez założeń arbitralnie ustalających jej postać, na przykład jako liniową lub logarytmiczną.

### 4.3. Przykładowe zastosowania do badań systemowych

Uniwersalny charakter estymatorów jądrowych umożliwia ich szerokie stosowanie w różnorodnych problemach współczesnej nauki i praktyki, ze szczególnym uwzględnieniem zagadnień interdyscyplinarnych badań systemowych. Proces ten jest możliwy dzięki dzisiejszej powszechności i coraz większym możliwościom nowoczesnych systemów komputerowych, a także zautomatyzowaniu procesów metrologicznych oraz pozyskiwania danych. Poniżej zostaną przedstawione **przykładowe zastosowania** estymatorów jądrowych do następujących zagadnień badań systemowych:

- rozpoznanie elementów nietypowych (zilustrowane możliwością stosowania do problemów archeologii),
- wyostrenie informacji nieprecyzyjnej (zilustrowane zagadnieniami współczesnej medycyny),
- identyfikacja parametryczna (zilustrowana zastosowaniem do sterowania automatycznego),
- określenie przestrzennego rozkładu popytu (zilustrowane zadaniem z zakresu telekomunikacji).

Najpierw zaprezentowane będą podstawowe elementy teorii decyzji, wykorzystywane wielokrotnie w dalszej części niniejszego podrozdziału.

### 4.3.1. Elementy teorii decyzji: reguły bayesowska i minimaksowa

Podstawowe zadanie **teorii decyzji** [2] polega na wyodrębnieniu ze zbioru wszystkich możliwych do podjęcia decyzji optymalnego elementu, jedynie na podstawie nieprecyzyjnych informacji o otaczającej rzeczywistości (naturze), w szczególności bez znajomości jej faktycznego stanu.

Niech więc dane będą:  $V$  – niepusty **zbiór możliwych stanów natury**,  $D$  – niepusty **zbiór potencjalnych decyzji**, a także **funkcja strat**  $l : V \times D \rightarrow \mathbb{R} \cup \{\pm\infty\}$ , której wartości  $l(v, d)$  interpretuje się jako straty wynikłe z podjęcia decyzji  $d$ , podczas gdy w rzeczywistości występuje stan  $v$ . Stratą minimaksową jest nazywane odwzorowanie  $l_m : D \rightarrow \mathbb{R} \cup \{\pm\infty\}$  zdefiniowane wzorem

$$l_m(d) = \sup_{v \in V} l(v, d). \quad (4.23)$$

Niech ponadto w zbiorze  $V$  określona będzie  $\sigma$ -algebra, a na niej miara  $\mathcal{M}$  reprezentująca rozkład nieokreśloności (w ujęciu probabilistycznym najczęściej charakteryzowana przez gęstość, a w ujęciu rozmytym – funkcję przynależności). Jeżeli dla każdego  $d \in D$  istnieje całka  $\int_V l(v, d) d\mathcal{M}(v)$ , to odwzorowanie  $l_b : D \rightarrow \mathbb{R} \cup \{\pm\infty\}$  dane jako

$$l_b(d) = \int_V l(v, d) d\mathcal{M}(v) \quad (4.24)$$

określa się mianem strat bayesowskich. Wtedy też każdy element  $d_m \in D$ , taki że

$$l_m(d_m) = \inf_{d \in D} l_m(d), \quad (4.25)$$

nosi nazwę **decyzji minimaksowej**. I analogicznie, dowolne  $d_b \in D$  spełniające warunek

$$l_b(d_b) = \inf_{d \in D} l_b(d) \quad (4.26)$$

jest **decyzją bayesowską**, a powyższe procedury są nazywane, odpowiednio, **regułą minimaksową** i **regułą bayesowską**. Druga z nich minimalizuje wartość oczekiwaną strat, podczas gdy pierwsza – największą stratę, jaka może nastąpić po podjęciu danej decyzji. Reguła bayesowska ma zatem charakter racjonalny, natomiast minimaksowa umożliwia asekurację przed wariantem skrajnie niekorzystnym.

### 4.3.2. Rozpoznanie elementów nietypowych

Zagadnienie rozpoznawania elementów nietypowych, dotychczas głównie wykorzystywane do analizy i eksploracji danych statystycznych, coraz powszechniej używane jest bezpośrednio do różnorodnych zadań praktycznych, takich jak wykrywanie uszkodzeń w systemach sterowania automatycznego, diagnostyka medyczna, archeologia i inne. Cel badań nierzadko stanowi tu podjęcie decyzji, czy w nadzorowanym systemie występują różnego rodzaju nieregularności, czy też jego stan można uznać za normalny. Innym przykładem jest stwierdzenie, który z pozyskanych obiektów archeologicznych nie pochodzi z okresu lub kultury, do której przynależą pozostałe.

W praktyce szczególnie dobrze nadającymi się do rozwiązywania problemów o powyższym charakterze są procedury oparte na **testowaniu hipotez statystycznych**. Zadanie często polega tu na identyfikacji tak zwanego **elementu odosobnionego** [1], czyli elementu w znacznym stopniu różniącego się od ogółu populacji, charakteryzowanej za pomocą próby losowej. Przy rozwiązywaniu tego typu zagadnień dominują tak zwane **testy istotności**, w których zakłada się, że prawdopodobieństwo błędu pierwszego rodzaju jest równe przyjętemu poziomowi istotności, a prawdopodobieństwo błędu drugiego rodzaju pozostaje tu niekontrolowane. W konsekwencji nie jest możliwe przyjęcie testowanej hipotezy, a jedynie jej odrzucenie. Jednakże już sam ewentualny brak decyzji o przyjęciu lub odrzuceniu hipotezy, implikujący niepodjęcie jakichkolwiek działań, okazuje się całkowicie wystarczający w wielu praktycznych zagadnieniach.

W rozważanym tu problemie rozpoznania elementu nietypowego zakłada się, że każdy element populacji jest scharakteryzowany za pomocą wartości  $n$  zmiennych ciągłych oraz  $k$  binarnych. Zmienne te powinny być tak dobrane, aby zbiory elementów typowych i nietypowych były jak najbardziej zróżnicowane w sensie zakresów wartości, jakie przyjmują poszczególne współrzędne, oraz przede wszystkim relacji między wartościami współrzędnych. Szczegółowa postać owych zbiorów i relacji nie musi być znana a priori – jej identyfikacja jest integralną częścią proponowanej tu procedury. Poniżej będzie przedstawiony test istotności przeznaczony do rozpoznawania elementów nietypowych na podstawie tak określonych wielkości, traktowanych jako  $(n + k)$ -wymiarowa zmienna losowa  $Z \equiv [X, Y]^T$ , zgodnie z oznaczeniami wprowadzonymi w punkcie 4.2.7.

Niech zatem  $\alpha \in (0, 1)$  stanowi przyjęty poziom istotności. Uzyskana ze zmiennej  $Z$  próba losowa (4.18) uznawana będzie jako wzorcowa, złożona z elementów typowych. Niech również dany będzie  $(n + k)$ -wymiarowy wektor  $\tilde{z} = [\tilde{x}, \tilde{y}]^T$ , który podlega sprawdzeniu, czy jest elementem nietypowym. Testuje się **hipotezę** stanowiącą, że  $\tilde{z}$  taki nie jest, przeciw hipotezie alternatywnej, iż należy go jednak uznać za nietypowy. Stosowana w tym celu **statystyka** określona będzie wzorem

$$S = \hat{f}(\tilde{z}), \quad (4.27)$$

gdzie  $\hat{f}$  oznacza estymator jądrowy gęstości wyznaczony na podstawie próby (4.18) traktowanej jako wzorcowa. O przyjęciu hipotezy alternatywnej stanowią zatem małe wartości statystyki  $S$ , co implikuje lewostronną postać **zbioru kry-**

tycznego

$$A = (-\infty, a], \quad (4.28)$$

przy czym **wartość krytyczna**  $a$  jest dana jako estymator kwantyla rzędu równego założonemu poziomowi istotności, zmiennej losowej  $U \equiv \hat{f} \circ Z$ , czyli jego wartość może być wyznaczona na podstawie próby

$$u_1 = \hat{f} \left( \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \right), u_2 = \hat{f} \left( \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \right), \dots, u_m = \hat{f} \left( \begin{bmatrix} x_m \\ y_m \end{bmatrix} \right) \quad (4.29)$$

z wykorzystaniem zależności (4.22). Dzięki przyjętej postaci statystyki (4.27) i własnościom estymatorów jądrowych tak skonstruowany test pozwala na identyfikację elementów nietypowych, różniących się od wzorcowych nie tylko zakresem wartości poszczególnych współrzędnych, ale nade wszystko złożonymi relacjami między tymi wartościami.

Własności estymatorów jądrowych umożliwiają uogólnienie powyższej, podstawowej postaci testu, przez zaprojektowanie dodatkowych funkcji zwiększających jego aplikacyjne możliwości: wersji lokalnej i warunkowej, zagadnienia rozmytości testowanego elementu, a także autokorekcji i adaptacji.

**Wersja lokalna** testu ma zastosowanie w przypadku używania jądra radialnego z transformacją liniową (4.10). Elementy macierzy  $R$  są wtedy definiowane jako

$$R_{k_1, k_2} = \frac{\sum_{i=1}^m \left[ x_{i, k_1} - \frac{1}{m} \left( \sum_{j=1}^m x_{j, k_1} \right) \right] \left[ x_{i, k_2} - \frac{1}{m} \left( \sum_{j=1}^m x_{j, k_2} \right) \right] K \left( \frac{x_i - \tilde{x}}{h} \right)}{\sum_{i=1}^m K \left( \frac{x_i - \tilde{x}}{h} \right)} \quad (4.30)$$

dla  $k_1, k_2 \in \{1, 2, \dots, n\}$ ,

gdzie  $K$  jest radialnym jądrem Epanecznikowa (tzn. danym wzorem (4.4) z podstawieniem (4.3)),  $h = \max_{i=1, 2, \dots, m} (|\tilde{x} - x_i|)$ , natomiast  $x_{i, k}$  i  $x_{j, k}$  oznaczają  $k$ -tą współrzędną, odpowiednio,  $i$ -tego i  $j$ -tego elementu próby losowej. Porównując powyższy wzór z definicją macierzy kowariancji, widać, że dzięki wprowadzeniu do formuły (4.30) jądra  $K$  odpowiednio preferowany jest wpływ elementów próby losowej bliższych testowanemu elementowi  $\tilde{x}$ .

W wielu zagadnieniach nauki i badań stosowanych występują **wielkości warunkujące** rozważany proces. Znajomość aktualnych ich wartości pozwala istotnie uściślić wiedzę o tym procesie. Typowym przykładem wielkości warunkującej w obiektach technicznych jest aktualna temperatura. Niech zatem zmienna losowa  $W$  o współrzędnych ciągłych i binarnych reprezentuje wielkości warunkujące. Gęstość łącznego rozkładu zestawienia  $[Z, W]^T$  może być wyznaczona z użyciem estymatora jądrowego. W danym zastosowaniu znajomość wielkości warunkującej, czyli wartości konkretnej realizacji  $W(\omega)$ , umożliwia dokonanie obliczeń na odpowiednim „przekroju” estymatora jądrowego, co pozwala znacząco uściślić wiedzę o badanym procesie.

W niektórych problemach aplikacyjnych poszczególne współrzędne testowanego elementu powinny być reprezentowane przez **liczby rozmyte**, pomimo

że próba (4.18) składa się z elementów „ostrych”. Próba ta jest bowiem złożona z elementów uznanych za wzorcowe i w tej sytuacji są one często określone precyzyjniej niż element testowany. Przykładowo, na próbę (4.18) mogą się składać wartości, jakie zostały zmierzone w przeszłości, podczas gdy element testowany dotyczy przyszłych, nieprecyzyjnie określonych wartości i jako taki w formie rozmytej lepiej oddaje modelowaną rzeczywistość. W przypadku gdy dowolna współrzędna elementu testowanego jest liczbą rozmytą, wtedy taka też staje się wartość statystyki (4.27). Jeżeli przyjąć konkretną postać tych liczb, na przykład trapezową lub typu L-R, to dzięki własnościom estymatorów jądrowych możliwe jest analityczne wyznaczenie funkcji przynależności wartości statystyki dla rozmytego elementu testowanego. O wyniku testu decyduje ostatecznie końcowa procedura oparta na **teorii preferencji rozmytych** [5], przy bezpośrednim zastosowaniu bayesowskiej reguły decyzyjnej (punkt 4.3.1).

Wynik prezentowanego testu może być błędny z dwóch powodów: element uznany za odosobniony taki nie był (błąd pierwszego rodzaju) lub element nieuznany za odosobniony taki był (błąd drugiego rodzaju). Jeżeli w rozpatrywanym zagadnieniu po pewnym czasie dysponuje się informacją, czy badany element był rzeczywiście odosobniony, czy też nie, to do konstrukcji statystyki (4.27) można użyć postaci uogólnionej (4.20) estymatora jądrowego, początkowo przyjmując  $w_i \equiv 1$ , i przez odpowiednią zmianę wartości tych współczynników modyfikować znaczenie poszczególnych jąder, stosownie do ich wpływu na błędny wynik. Koncepcja ta może być zrealizowana z użyciem sztucznych sieci neuronowych, przy wykorzystaniu metody wrażliwościowej. Uzyskuje się dzięki temu możliwość **autokorekcji** działania testu oraz jego **adaptacji**, postępującej w miarę zmian uwarunkowań badanego zjawiska.

Prezentowana metodyka rozpoznawania elementów nietypowych ma charakter uniwersalny, a jej użyteczność jest weryfikowana w problemie z zakresu archeologii. Przedmiot badań stanowi stwierdzenie, czy pochodzący z wykopalisk prześlik został wyprodukowany w rejonie miejscowości Owrucz (Rosja). Poszczególnymi współrzędnymi zmiennej losowej  $X$  są wymiary, waga (współrzędne ciągłe) i inne cechy (ujęte w sposób binarny) poszczególnych prześlików. Jako próbę wzorcową przyjęto typowe, zweryfikowane klasycznymi metodami archeologicznymi obiekty z Owrucza, z dokładnie określonymi wielkościami stanowiącymi poszczególne współrzędne zmiennej  $X$ . Testowane elementy są często zniszczone lub niestarannie wykonane, a zatem niektóre ich współrzędne należy traktować jako liczby rozmyte.

Typ i kształt jądra  $K$  użytego do konstrukcji statystyki (4.27) zależny był od wersji testu. I tak, w wersji podstawowej stosowano jądro radialne normalne, czyli na podstawie zależności (4.4) z podstawieniem

$$\mathcal{K}(x) = (2\pi)^{-1/2} \exp\left(-\frac{1}{2}x^2\right). \quad (4.31)$$

Podobnie postępowano przy wersji lokalnej testu. W wersji warunkowej i w przypadku rozmytego elementu testowanego wykorzystane było produktowe jądro normalne, to znaczy oparte na wzorach (4.5) i (4.31). Wprowadzone zostały wszystkie awizowane dodatkowe procedury, czyli modyfikacja parametru wy-

gładzania (4.9), w przypadku stosowania jądra radialnego – transformacja liniowa (4.10), ograniczenie nośnika poszczególnych współrzędnych (4.13), a także współrzędne binarne (4.19) zarówno dla zmiennej podstawowej  $Z$ , jak i warunkowej  $W$ . Do estymacji kwantyla (4.22) użyto uogólnionego jądra Cauchy'ego

$$\mathcal{K}(x) = \frac{2}{\pi} \frac{1}{(1+x^2)^2}, \quad (4.32)$$

z modyfikacją parametru wygładzania (4.9) i ograniczeniem nośnika (4.13) do zbioru liczb nieujemnych.

Dane archeologiczne pozyskano dzięki współpracy z Instytutem Archeologii i Etnologii PAN.

Niniejszy punkt ujmuje wyniki przedstawione w monografii [10], a także badania prowadzone wraz z Cyprianem Prochotem w ramach Jego pracy doktorskiej [29] – w tym zakresie tekst prezentuje materiał wspólnych publikacji [23, 24].

### 4.3.3. Wyostrzanie informacji nieprecyzyjnej

Rozwój współczesnych technik informacyjnych umożliwia ciągły postęp w zakresie wnioskowania, a w konsekwencji podejmowania działań na podstawie **informacji nieprecyzyjnej**. Konieczność operowania taką informacją wynika najczęściej ze złożoności badanych systemów oraz intuicyjnego charakteru ludzkich wrażeń. Jednakże praktyczne zagadnienia wymagają podjęcia na tej podstawie konkretnej decyzji, często łączącej się z koniecznością jednoznacznego wskazania ściśle określonej wartości liczbowej. Obecnie wnioskowanie takie – właściwe dotychczas jedynie człowiekowi – jest coraz powszechniej przeprowadzane, lub przynajmniej wspomagane, przez systemy informatyczne. W bardziej złożonych zagadnieniach wiąże się to nierzadko z dodatkowym uwzględnieniem aktualnych wartości czynników warunkujących, umożliwiającymi uściślenie posiadanych informacji, co klasyczny proces wnioskowania często uczyniłoby zbyt skomplikowanym.

Jako prosta **ilustracja zadania**, do którego stosowana może być proponowana poniżej procedura, rozważony będzie klasyczny problem medyczny określenia „odpowiedniej dawki leku”. Ze względu na charakter problemów większość ocen pozyskiwanych w podobnych zagadnieniach ma charakter uznaniowy, a z uwagi na specyfikę środowiska medycznego są one nierzadko formułowane w sposób werbalny, na przykład „dobre”, „złe”, „wątpliwe”. Istnieje też wiele zmiennych warunkujących dany problem, jak waga, wiek i płeć pacjenta, o ciągłym i binarnym charakterze. Co więcej, nieprawidłowa diagnoza, oprócz skutków ekonomicznych związanych z kosztami leczenia, często implikuje z jednej strony ewentualność powikłań, a z drugiej niebezpieczeństwo nieskuteczności terapii. Wymusza to konieczność uwzględnienia odmiennych skutków przeszacowania i niedoszacowania podjętej decyzji, proporcjonalnych do wartości owego przeszacowania lub niedoszacowania, ale o różnych współczynnikach proporcjonalności. Poniżej przedstawiony będzie algorytm wyostrzania informacji nieprecyzyjnej, czyniący zadość powyższym uwarunkowaniom.

Przyjmuje się, że informacja nieprecyzyjna dotyczy wielkości jednowymiarowej (rzeczywistej), przy czym miara tej informacji jest scharakteryzowana przez gęstość  $\varphi : \mathbb{R} \rightarrow [0, \infty)$  (przykładowo w ujęciu probabilistycznym jej odpowiednikiem jest gęstość rozkładu, natomiast w ujęciu rozmytym – funkcja przynależności), przy czym  $\varphi \in L^1$ . Niech ponadto dana będzie  $(n+k)$ -wymiarowa zmienna losowa  $Z \equiv [X, Y]^T$ , będąca zestawieniem  $n$ -wymiarowej ciągłej zmiennej  $X$ , której rozkład ma gęstość, oraz  $k$ -wymiarowej zmiennej binarnej  $Y$ . Zmienna  $Z$  reprezentuje czynnik warunkujący.

Jeżeli rozważana jest przestrzeń wektorowa  $\mathbb{R}^{n+k+1}$ , której pierwsza współrzędna reprezentuje zmienną niezależną funkcji  $\varphi$ , a pozostałe – kolejne współrzędne zmiennej losowej  $Z$ , to w podprzestrzeni złożonej z pierwszej współrzędnej dana jest miara charakteryzowana przez gęstość  $\varphi$ , a w podprzestrzeni złożonej z pozostałych  $n+k$  współrzędnych – miara będąca rozkładem zmiennej losowej  $Z$ . Tak więc w całej przestrzeni  $\mathbb{R}^{n+k+1}$  określona jest miara produktowa, której rozkład można oszacować za pomocą estymatora jądrowego  $\hat{f}$ . Jeżeli gęstość  $\varphi$  wyznacza się na podstawie werbalnych ocen (reprezentowanych przez współczynniki  $w_i$ ) dotyczących wartości  $v_i$  otrzymanych przy różnych wartościach zmiennej warunkującej  $z_i = [x_i, y_i]^T$ , to powinna być stosowana uogólniona postać estymatora jądrowego (4.20).

W prezentowanej wcześniej ilustracji problemu zostało stwierdzone, że niedoszacowanie i przeszacowanie podjętej decyzji skutkują stratami proporcjonalnymi do wartości owego przeszacowania lub niedoszacowania, ale o różnych współczynnikach proporcjonalności. Bayesowska reguła decyzyjna dostarcza stosownego aparatu do rozwiązania takiego zadania, przy czym  $D = \mathbb{R}$ , a **funkcja strat**  $l$  przyjmuje postać

$$l(\hat{v}, v) = \begin{cases} -p(\hat{v} - v) & \text{dla } \hat{v} - v \leq 0, \\ q(\hat{v} - v) & \text{dla } \hat{v} - v \geq 0, \end{cases} \quad (4.33)$$

gdzie  $p, q > 0$  (przy czym współczynniki te mogą być różne), natomiast  $\hat{v}$  oznacza szukaną wartość decyzji bayesowskiej, pełniącą tu rolę determinizatora.

Niech będzie ustalony czynnik losowy, a zatem także i wartość  $Z(\omega) = [X(\omega), Y(\omega)]^T$ . Determinizator jest wtedy dany jako rozwiązanie następującego równania z argumentem  $\hat{v}$ :

$$\int_{-\infty}^{\hat{v}} \hat{f}(v, Z(\omega)) dv = \frac{p}{p+q} \int_{-\infty}^{\infty} \hat{f}(v, Z(\omega)) dv, \quad (4.34)$$

gdzie  $\hat{f}$  oznacza wprowadzony powyżej estymator jądrowy. Oznaczając jako  $I : \mathbb{R}^{n+1} \rightarrow [0, \infty)$  funkcję pierwotną jądra  $K$  względem pierwszej współrzędnej, czyli

$$I \left( \begin{bmatrix} v \\ x \end{bmatrix} \right) = \int_{-\infty}^v K \left( \begin{bmatrix} u \\ x \end{bmatrix} \right) du \quad \text{dla dowolnie ustalonego } x, \quad (4.35)$$

równanie (4.34) przyjmuje postać



$$\begin{aligned} \sum_{i=1}^m w_i I \left( \frac{\begin{bmatrix} \hat{v} \\ X(\omega) \end{bmatrix} - \begin{bmatrix} v_i \\ x_i \end{bmatrix}}{h} \right) L(Y(\omega), y_i) = \\ = \frac{p}{p+q} \sum_{i=1}^m w_i \int_{-\infty}^{\infty} K \left( \frac{\begin{bmatrix} v \\ X(\omega) \end{bmatrix} - \begin{bmatrix} v_i \\ x_i \end{bmatrix}}{h} \right) L(Y(\omega), y_i) dv, \end{aligned} \quad (4.36)$$

gdzie  $w_i$  oznaczają oceny ostrych wartości  $v_i$  uzyskanych, odpowiednio, dla ciągłych  $x_i$  i binarnych  $y_i$  wartości czynnika warunkującego. Rozwiązanie powyższego równania stanowi ostatecznie poszukiwany determinizator przy ustalonym czynniku warunkującym. Jeżeli jądro  $K$  ma dodatnie wartości, to powyższe rozwiązanie istnieje i jest jedyne. W praktyce może być ono efektywnie obliczone za pomocą iteracyjnej metody Newtona, gdyż zgodnie ze wzorem (4.35) pochodna lewej strony jest prosta do wyznaczenia.

W przeprowadzonych badaniach użyto radialnego jądra opartego na uogólnionym jądrze Cauchy'ego

$$\mathcal{K}(x) = \frac{1}{A_n} \frac{1}{(1 + x^T x)^{a_n}}, \quad (4.37)$$

przy  $a_n = \left\lceil \frac{n+5}{2} \right\rceil$  oraz  $A_n = (n+1)V_{n+1} \int_0^\infty \frac{r^n}{(1+r^2)^{a_n}} dr$ , gdzie  $[b]$  oznacza część całkowitą liczby  $b \in \mathbb{R}$ , a  $V_{n+1}$  jest objętością  $(n+1)$ -wymiarowej kuli jednostkowej. Zastosowane zostały wszystkie zaprezentowane dodatkowe procedury, czyli modyfikacja parametru wygładzania (4.9), transformacja liniowa (4.10), ograniczenie nośnika (4.13), a także współrzędne binarne (4.19).

**Podsumowując:** w niniejszym punkcie została przedstawiona procedura bayesowskiego wyostrzania informacji nieprecyzyjnej w obecności losowych czynników warunkujących. Chociaż możliwości jej wykorzystania są uniwersalne, to zagadnienie zilustrowano w aspekcie współczesnych zastosowań medycznych na modelowym zadaniu ustalenia „prawidłowej dawki lekarstwa”. Informacja posiadana w tym zakresie jest nieprecyzyjna, reprezentowana przez gęstość  $\varphi$ , a uzyskiwane oceny mają charakter uznaniowo-werbalny. W omawianym przykładzie wymaga się, aby dla pewnej zbiorowości pacjentów, indeksowanej przez  $i = 1, 2, \dots, m$ , uzyskać dla zaordynowanych poszczególnym osobom dawek lekarstwa  $v_i$  werbalną ocenę ich skuteczności „bardzo dobra”, „dobra”, „prawidłowa”, „zła” itd. – ocenom tym przyporządkować można wartości współczynników  $w_i$ . „Prawidłowa dawka lekarstwa” zależy silnie od wieku, masy ciała (zmienne ciągłe  $X_1$  i  $X_2$ ) oraz płci (zmienna binarna  $Y$ ), które dla poszczególnych pacjentów wynoszą  $[x_{1i}, x_{2i}, y_i]^T$ . Jeżeli rozważana jest cała populacja, to czynniki te są losowe, reprezentowane przez 3-wymiarową zmienną losową  $Z \equiv [X_1, X_2, Y]^T$ . Miara produktowa informacji nieprecyzyjnej i rozkładu zmiennej  $Z$  została scharakteryzowana za pomocą statystycznego estymatora jądrowego (4.20). Do jego wyznaczenia korzysta się ze zbioru deterministycznych danych typu

$$\begin{bmatrix} (v_i, w_i) \\ x_{1i} \\ x_{2i} \\ y_i \end{bmatrix} \quad (4.38)$$

uzyskanych na podstawie  $m$  przypadków klinicznych.

Jeżeli rozważany jest konkretny pacjent, dla którego ustalona ma być „prawidłowa dawka lekarstwa”, to wartość warunkującej zmiennej losowej  $Z(\omega)$  staje się ustalona. Warto zauważyć, że w zbiorze zebranych wartości warunkujących  $[x_{1i}, x_{2i}, y_i]^T$  może w ogóle nie być wartości równej realizacji  $Z(\omega) = [X(\omega), Y(\omega)]^T$  lub zebrana niżej ich ilość absolutnie może nie wystarczać do odpowiedniego wnioskowania. Tymczasem własności estymatorów jądrowych pozwalają odpowiednio uśrednić posiadane dane. W praktyce pacjentowi należy podać konkretną dawkę lekarstwa. Niedoścadowanie dawki ma najczęściej zdecydowanie odmienny wpływ na ogólne koszty leczenia niż przeszacowanie. Stąd zadanie wyostrzenia informacji nieprecyzyjnej powinno być prowadzone przy niesymetrycznej funkcji strat (4.31).

Przedstawiona procedura została pozytywnie zweryfikowana na podstawie danych pozyskanych z Samodzielnego Publicznego Szpitala Klinicznego Nr 1 w Lublinie oraz wyników eksperymentalnych otrzymanych w Akademii Medycznej w Lublinie.

Niniejszy punkt ujmuje wyniki badań prowadzonych wraz z Małgorzatą Charytanowicz, z Akademii Medycznej w Lublinie, w ramach Jej pracy doktorskiej [3]. Powyższy tekst zawiera fragmenty tej pracy oraz wspólnych publikacji [19, 20].

#### 4.3.4. Identyfikacja parametryczna

Jednym z podstawowych zagadnień badań systemowych jest **identyfikacja parametryczna**, czyli określenie wartości parametrów występujących w modelu. W typowym zadaniu praktycznym dysponuje się  $m$  niezależnymi pomiarami danego parametru, które są jednak obciążone błędami różnego pochodzenia. Na tej podstawie należy określić wartość, która możliwie najlepiej – z punktu widzenia całości problemu – charakteryzowałaby zjawiska reprezentowane przez ten parametr. Typowe procedury estymacji, oparte na metodach najmniejszych kwadratów i największej wiarygodności, są używane najczęściej ze względu na powszechną znajomość i dostępność w literaturze, ale nie pozwalają one uwzględnić zróżnicowanych skutków błędów estymacji.

Zagadnienie to zostanie zilustrowane na przykładzie **sterowania optymalnego**. Będący podstawą takiego zadania wskaźnik jakości może być bowiem użyty do określenia nie tylko jakości sterowania, lecz także procedury identyfikacji parametrów modelu. Niech zatem dany będzie układ, którego dynamika jest opisana za pomocą równania różniczkowego

$$\dot{x}(t) = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ \lambda \end{bmatrix} u(t), \quad (4.39)$$

przy czym  $\lambda \in \mathbb{R} \setminus \{0\}$ . Jeżeli regulator optymalny z kwadratowym wskaźnikiem jakości został wyznaczony na podstawie wartości  $\lambda \in \mathbb{R} \setminus \{0\}$  niekoniecznie równej parametrowi  $\lambda$  występującemu w obiekcie, to otrzymany wykres tego wskaźnika z dużą dokładnością można aproksymować funkcją kwadratową o różnych współczynnikach dla błędów ujemnych i dodatnich. Traktując uzyskane wartości badanego parametru jako realizacje zmiennej losowej z rozkładem o gęstości  $f$ , wyznaczyć można za pomocą reguły bayesowskiej wartość estymatora optymalnego w sensie minimum wartości oczekiwanej kwadratowego wskaźnika jakości sterowania. Jeżeli rozkład powyższej zmiennej losowej estymuje się za pomocą estymatorów jądrowych, to uniezależnia to metodę od typu rozkładu, a uzyskany algorytm jest dogodny dla procedur numerycznych.

Najpierw będzie rozpatrzony podstawowy przypadek, gdy rozważany jest pojedynczy parametr. Jako zbiór decyzji dopuszczalnych przyjmuje się wówczas  $D = \mathbb{R}$ , a funkcja strat zostanie określona w postaci niesymetrycznej kwadratowej

$$l(\hat{v}, v) = \begin{cases} p(\hat{v} - v)^2 & \text{dla } \hat{v} - v \leq 0, \\ q(\hat{v} - v)^2 & \text{dla } \hat{v} - v \geq 0, \end{cases} \quad (4.40)$$

gdzie  $p, q > 0$  (przy czym współczynniki te mogą być różne), natomiast  $\hat{v}$  oznacza szukaną wartość decyzji bayesowskiej, pełniącą tu rolę optymalnego estymatora. Jest on wtedy dany jako rozwiązanie następującego równania z argumentem  $\hat{v}$ :

$$(p - q) \int_{-\infty}^{\hat{v}} (\hat{v} - v) f(v) dv = p \int_{-\infty}^{\infty} (Z - z) f(v) dv. \quad (4.41)$$

Rozwiązanie powyższego kryterium nie jest w ogólnym przypadku zadaniem łatwym. Jeżeli jednak do oszacowania gęstości  $f$  są stosowane estymatory jądrowe, to dzięki odpowiedniemu doborowi postaci jądra można otrzymać efektywny algorytm numeryczny. Niech zatem dostępnych będzie  $m$  pomiarów badanego parametru, traktowanych jako próba losowa (4.1). Dla wybranego jądra  $K$  można zdefiniować następujące rzeczywiste odwzorowania dane wzorami:

$$I(x) = \int_{-\infty}^x K(y) dy, \quad (4.42)$$

$$J(x) = \int_{-\infty}^x y K(y) dy \quad (4.43)$$

oraz (w przypadku podstawowej postaci estymatora jądrowego (4.2)):

$$\bar{I}(x) = \frac{1}{mh} \sum_{i=1}^m I\left(\frac{x - x_i}{h}\right), \quad (4.44)$$

$$\bar{J}(x) = \frac{1}{mh} \sum_{i=1}^m J\left(\frac{x - x_i}{h}\right). \quad (4.45)$$

Wtedy kryterium (4.40) przyjmuje postać równania

$$(p - q)[\hat{v}\bar{I}(\hat{v}) - \bar{J}(\hat{v})] - p\hat{v} = -p\frac{1}{m}\sum_{i=1}^m x_i. \quad (4.46)$$

Jeżeli jądro  $K$  ma dodatnie wartości, to powyższe rozwiązanie istnieje i jest jedyne. Oznaczając jego lewą i prawą stronę, odpowiednio, przez

$$L(\hat{v}) = (p - q)[\hat{v}\bar{J}(\hat{v}) - \bar{I}(\hat{v})] - p\hat{v}, \quad (4.47)$$

$$P = -p\frac{1}{m}\sum_{i=1}^m x_i \quad (4.48)$$

oraz wyznaczając pochodną funkcji (4.47):

$$L'(\hat{v}) = (p - q)\bar{J}(\hat{v}) - p, \quad (4.49)$$

można wtedy – korzystając z iteracyjnej metody Newtona – efektywnie wyznaczyć rozwiązanie kryterium (4.40) jako granicę ciągu  $\{\hat{v}^{(i)}\}_{i=1}^{\infty}$  danego wzorami

$$v^{(0)} = \frac{1}{m}\sum_{i=1}^m x_i, \quad (4.50)$$

$$v^{(k+1)} = v^{(k)} + \frac{P - L(v^{(k)})}{L'(v^{(k)})} \quad \text{dla } k = 0, 1, \dots \quad (4.51)$$

Własności estymatorów jądrowych umożliwiają wprowadzenie uogólnień powyższej metody dla przypadku wielowymiarowego (wektor parametrów), warunkowego (zależność od czynników warunkujących) oraz wielomianowego (funkcja strat stopnia większego niż kwadratowy).

I tak, w przypadku wielowymiarowym, czyli zagadnieniu **estymacji wektora parametrów**, możliwe jest powtórzenie proponowanej procedury wobec wielowymiarowej zmiennej losowej. Przykładowo, jeśli dla dwuwymiarowego wektora parametrów  $v = [v_1, v_2]^T$  przyjąć funkcję strat w następującej postaci niesymetrycznej i kwadratowej:

$$l\left(\begin{bmatrix} \hat{v}_1 \\ \hat{v}_2 \end{bmatrix}, \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}\right) = \begin{cases} c_l(\hat{v}_1 - v_1)^2 + c_{ld}(\hat{v}_1 - v_1)(\hat{v}_2 - v_2) + c_d(\hat{v}_2 - v_2)^2, \\ \quad \text{gdy } \hat{v}_1 - v_1 \leq 0 \quad \text{i} \quad \hat{v}_2 - v_2 \leq 0, \\ c_p(\hat{v}_1 - v_1)^2 + c_{pd}(\hat{v}_1 - v_1)(\hat{v}_2 - v_2) + c_d(\hat{v}_2 - v_2)^2, \\ \quad \text{gdy } \hat{v}_1 - v_1 \geq 0 \quad \text{i} \quad \hat{v}_2 - v_2 \leq 0, \\ c_l(\hat{v}_1 - v_1)^2 + c_{lg}(\hat{v}_1 - v_1)(\hat{v}_2 - v_2) + c_g(\hat{v}_2 - v_2)^2, \\ \quad \text{gdy } \hat{v}_1 - v_1 \leq 0 \quad \text{i} \quad \hat{v}_2 - v_2 \geq 0, \\ c_p(\hat{v}_1 - v_1)^2 + c_{pg}(\hat{v}_1 - v_1)(\hat{v}_2 - v_2) + c_g(\hat{v}_2 - v_2)^2, \\ \quad \text{gdy } \hat{v}_1 - v_1 \geq 0 \quad \text{i} \quad \hat{v}_2 - v_2 \geq 0, \end{cases} \quad (4.52)$$

przy czym  $c_l, c_p, c_g, c_d > 0$ ,  $c_{ld}, c_{pg} \geq 0$  oraz  $c_{pd}, c_{lg} \leq 0$ , to kryterium (4.40) przyjmuje tu postać dwóch równań definiujących decyzje bayesowskie  $\hat{v}_1$  oraz  $\hat{v}_2$  o podobnej, aczkolwiek nieco bardziej złożonej postaci.

W przypadku gdy rozważany parametr jest silnie zależny od czynnika warunkującego, proponowana procedura może być uogólniona w celu uwzględnienia jego wpływu. Niech zatem czynnik warunkujący dany będzie w postaci zmiennej losowej  $Z$ , stanowiącej zestawienie  $n$ -wymiarowej ciągłej zmiennej, której rozkład ma gęstość, a także  $k$ -wymiarowej zmiennej binarnej. Statystyczne estymatory jądrowe mogą być użyte do wyestymowania gęstości rozkładu łącznego. W trakcie stosowania procedury, po ustaleniu konkretnej wartości czynnika warunkującego  $Z(\omega)$ , kryterium (4.41) może być stosowane „na przekroju” definiowanym przez ową wartość. W przypadku wystąpienia istotnego uwarunkowania badanego obiektu od czynników reprezentowanych przez zmienną  $Z$ , ujęcie takie może znacząco polepszyć jakość uzyskanych wyników.

Postać niesymetrycznej kwadratowej funkcji strat (4.40) może być uogólniona do formuły wielomianowej:

$$l(\hat{v}, v) = \begin{cases} (-1)^k p(\hat{v} - v)^k & \text{dla } \hat{v} - v \leq 0, \\ q(\hat{v} - v)^k & \text{dla } \hat{v} - v \geq 0, \end{cases} \quad (4.53)$$

przy czym  $p, q > 0$  oraz  $k = 2, 3, \dots$ . Przykładowo, dla  $k = 3$  kryterium analogiczne do równania (4.41) przyjmuje postać

$$(p + q) \int_{-\infty}^{\hat{v}} (\hat{v} - v)^2 f(v) dv = q \int_{-\infty}^{\infty} (\hat{v} - v)^2 f(v) dv. \quad (4.54)$$

Efektywność przedstawionej tu procedury w dużym stopniu zależy od wyboru postaci jądra  $K$ , w szczególności od możliwości wyrażenia formuł analitycznych funkcji  $I$  oraz  $J$ . W przypadku podstawowym, gdy rozważany był pojedynczy parametr, zostało użyte uogólnione jądro Cauchy'ego

$$K(x) = \frac{2}{\pi} \frac{1}{(1 + x^2)^2}. \quad (4.55)$$

Wtedy

$$I(x) = \frac{1}{\pi} \left[ \operatorname{arctg}(x) + \frac{x}{(1 + x^2)} \right] + \frac{1}{2}, \quad (4.56)$$

$$J(x) = \frac{-1}{\pi(1 + x^2)}. \quad (4.57)$$

W przypadku wielowymiarowym i warunkowym stosowane jest jądro produktowe (4.5) z podstawieniem zależności (4.55), aczkolwiek w przypadku wielomianowym wymaga ona uściślenia do postaci

$$\mathcal{K}(x) = \frac{1}{A_k} \frac{1}{(1 + x^2)^{a_k}}, \quad (4.58)$$

przy  $a_k = \left[ \frac{(k + 3)}{2} \right]$  oraz  $A_k = 2 \int_0^{\infty} \frac{1}{(1 + r^2)^{a_k}} dr$ , gdzie  $[b]$  stanowi część całkowitą liczby  $b \in \mathbb{R}$ .

Przedstawiona procedura została zweryfikowana w Zakładzie Automatyki (E-32) Politechniki Krakowskiej.

Niniejszy punkt ujmuje wyniki badań prowadzonych wraz z Aleksandrem Mazgajem, z tegoż Zakładu, w ramach Jego pracy doktorskiej [28]. Powyższy tekst zawiera fragmenty tej pracy oraz wspólnych publikacji [21, 22].

#### 4.3.5. Określenie przestrzennego rozkładu popytu

**System LMDS** (*Local Multipoint Distribution System*) jest coraz powszechniej stosowany przez operatorów telekomunikacyjnych do bezprzewodowej szerokopasmowej transmisji danych, co w dużym stopniu wynika z gwałtownie narastającego popytu na dostęp do Internetu. System ten umożliwia połączenie węzła sieci operatorskiej z odbiorcami usług, bez konieczności budowy kosztownej infrastruktury kablowej. Dane są przekazywane między rozlokowanymi na terenie aglomeracji miejskiej stacjami bazowymi, które drogą radiową obsługują stałe połączenia z wieloma stacjami abonenckimi, znajdującymi się w zasięgu transmisji urządzeń nadawczo-odbiorczych poszczególnych stacji. Stacje abonenckie, instalowane najczęściej na dachach budynków, transmitują z kolei – za pomocą sieci lokalnych, na przykład kablowych – owe dane do użytkowników w obrębie danej lokalizacji. Istotnym czynnikiem, często decydującym o ekonomicznej zasadności stosowania tego typu systemu, jest wyznaczenie miejsc zainstalowania stacji bazowych w taki sposób, aby w granicach dostępnych środków inwestycyjnych osiągnąć największy zysk. W niniejszym punkcie zaprezentowano algorytm wyznaczania optymalnego układu stacji bazowych systemu LMDS.

**Przestrzenny rozkład popytu** na usługi transmisji danych w rozpatrywanym terenie ma – w dosłownym znaczeniu – charakter punktowy związany z poszczególnymi potencjalnymi klientami. Taki model jednak charakteryzuje się ograniczonym znaczeniem aplikacyjnym, gdyż praktycznie jest nieidentyfikowalny na obszarze aglomeracji miejskiej. Do oszacowania przestrzennego rozkładu popytu zostaną tutaj użyte statystyczne estymatory jądrowe. Występująca w ich definicji zmienna  $X$  jest zatem 2-wymiarowa, a jej poszczególne współrzędne reprezentują długość i szerokość geograficzną. Zastosowane zostało dwuwymiarowe uogólnione radialne jądro Cauchy'ego, czyli formuła (4.4) po podstawieniu (4.32), z modyfikacją parametru wygładzania (4.9) i po przyporządkowaniu poszczególnym jądom współczynników (4.20). I tak, dysponując bazą danych złożoną z  $m$  potencjalnych lokalizacji budynków abonenckich, przy czym każda z nich jest charakteryzowana przez jej położenie geograficzne  $x_i = [x_{i1}, x_{i2}]^T$  oraz współczynnik  $w_i$  reprezentujący potencjalny popyt na usługi transmisji danych przypisany tej lokalizacji, można otrzymać estymator jądrowy opisujący gęstość przestrzennego rozkładu popytu na rozważanym obszarze. Rozkład ten jest odpowiednio uciąglony dzięki własnościom statystycznych estymatorów jądrowych. Co więcej, z uwagi na uśredniające własności tego typu estymatorów możliwe staje się korzystanie z uproszczonej bazy danych zawierającej jedynie lokalizacje głównych budynków abonenckich i zawarcie w przypisanych im współczynnikach  $w_i$  popytu sąsiednich, pomniejszych obiektów. Znacznie upraszcza to najtrudniejszą i najbardziej kosztowną fazę rozważanej tu procedury wyznaczania optymalnego układu stacji bazowych systemu LMDS.

W praktyce nietrudno jest wskazać ograniczoną liczbę  $k$  miejsc, w których można ewentualnie dokonać montażu stacji bazowych, obejmujących na przykład wysokie budynki i maszty telekomunikacyjne. Dysponując estymatorem jądrowym  $\hat{f}$ , charakteryzującym przestrzenny rozkład popytu w rozważanym obszarze, określa się wskaźnik jakości  $j$ -tej lokalizacji ( $j = 1, 2, \dots, k$ ) w następującej postaci:

$$E_j = \int_{C_j} \hat{f}(x) dx, \quad (4.59)$$

gdzie  $C_j$  oznacza  $j$ -te koło o środku w  $x_j$  i dodatnim promieniu  $r_j$  reprezentującym zasięg urządzenia nadawczo-odbiorczego przypisanego tej lokalizacji. Dzięki odpowiedniej postaci jądra możliwe jest analityczne wyznaczenie tej wartości za pomocą wzoru

$$E_j = \frac{1}{2 \sum_{i=1}^m w_i} \sum_{i=1}^m w_i \left( \frac{r_j^2 - d_{ij}^2 - h^2 s_i^2}{\sqrt{r_j^4 + 2[h^2 s_i^2 - d_{ij}^2]r_j^2 + [h^2 s_i^2 + d_{ij}^2]^2}} + 1 \right), \quad (4.60)$$

przy czym  $d_{ij}$  oznacza odległość między  $j$ -tą lokalizacją a środkiem  $i$ -tego jądra, a ponieważ z punktu widzenia prowadzonych w dalszej części zagadnień optymalizacji dodatnia stała  $\frac{1}{2} \sum_{i=1}^m w_i$  może być pominięta, więc ostatecznie

$$E_j = \sum_{i=1}^m w_i \left( \frac{r_j^2 - d_{ij}^2 - h^2 s_i^2}{\sqrt{r_j^4 + 2[h^2 s_i^2 - d_{ij}^2]r_j^2 + [h^2 s_i^2 + d_{ij}^2]^2}} + 1 \right). \quad (4.61)$$

Na podstawie powyższej zależności można ustalić **wskaźnik jakości** dowolnego układu stacji bazowych z przypisanymi poszczególnym lokalizacjom typami urządzeń nadawczo-odbiorczych (lub ich brakiem, gdy w rozważanej opcji nie jest planowane zamontowanie takich urządzeń). Proponowana procedura umożliwia także uwzględnienie obszarów „zacienionych”, w których transmisja nie jest możliwa ze względu na nierównomierne ukształtowanie terenu oraz przeszkody (np. wysokie budynki), a także ograniczonej możliwości przesyłowej (przepustowości) urządzeń nadawczo-odbiorczych. W atrakcyjnych rejonach miast popyt objęty zasięgiem poszczególnych urządzeń może bowiem nie zostać zaspokojony z jednej stacji bazowej i część niezaspokojonego popytu powinna przejąć inna stacja bazowa będąca w zasięgu – optymalny podział jest wyznaczany na podstawie metod optymalizacji liniowej.

Dysponując skonstruowanym powyżej wskaźnikiem jakości ustalonego układu stacji bazowych, można wyznaczyć – korzystając z metod badań operacyjnych, w szczególności metody podziału i ograniczeń – optymalny ich układ w ramach posiadanych środków inwestycyjnych. Możliwe jest także poszerzenie problemu na zagadnienia planowania z kilkuletnim horyzontem czasowym, przy niestacjonarności uwarunkowań.

Wprowadzone wcześniej współczynniki  $w_i$  reprezentują zapotrzebowanie na usługi teletransmisyjne przypisane poszczególnym lokalizacjom stacji abonenckich. Ich identyfikacja odbywa się w praktyce na podstawie oceny eksperta, wyrażanej werbalnie, często jedynie na podstawie przesłanek intuicyjnych. Do opisu przewidywanego w stacji abonenckiej zapotrzebowania na usługi teletransmisyjne zostaną zatem użyte elementy **logiki rozmytej**. Zagadnienie rozmytego charakteru współczynników  $w_i$  wprowadzonych w definicji (4.20) zostanie szczegółowo omówione poniżej, gdyż zadanie takie może pojawić się w różnorodnych zastosowaniach estymatorów jądrowych do zagadnień badań systemowych. Należy tu wziąć pod uwagę specyfikę rozważanego zadania: znaczną ilość koniecznych do zidentyfikowania i późniejszej analizy liczb rozmytych (równą liczności próby  $m$ ), a także to, iż czasem niektóre współczynniki  $w_i$  mogą być „ostre” (nierozmyte).

Szczególnie dogodne w powyższej sytuacji są **liczby rozmyte typu L-R** [8], których funkcja przynależności jest określona w następującej postaci:

$$\mu_{(w_i, \alpha_i, \beta_i)}(x) = \begin{cases} L\left(\frac{w_i - x}{\alpha_i}\right) & \text{dla } x \leq w_i, \\ R\left(\frac{x - w_i}{\beta_i}\right) & \text{dla } x \geq w_i, \end{cases} \quad (4.62)$$

przy czym  $w_i \in \mathbb{R}$ ,  $\alpha_i, \beta_i > 0$ , a rzeczywiste funkcje  $L$  i  $R$  są symetryczne względem zera, przyjmują tam wartość 1 i są niemalejące na przedziale  $(-\infty, 0]$ . W tym układzie parametr  $w_i$  można zinterpretować jako wartość modalną, natomiast  $\alpha_i$  i  $\beta_i$  decydują, odpowiednio, o lewostronnym i prawostronnym skupieniu wokół tej wartości. Liczbę rozmytą  $\mathcal{A}$  typu L-R można zatem utożsamiać z trzema parametrami, co oznaczane będzie jako  $\mathcal{A} = (w, \alpha, \beta)$ , i w konsekwencji proces identyfikacji wymaga określenia jedynie tych, bliskich intuicyjnej interpretacji wartości. Działania algebraiczne na liczbach rozmytych typu L-R definiowane są następująco:

$$\mathcal{A} + \mathcal{B} = (w_{\mathcal{A}}, \alpha_{\mathcal{A}}, \beta_{\mathcal{A}}) + (w_{\mathcal{B}}, \alpha_{\mathcal{B}}, \beta_{\mathcal{B}}) = (w_{\mathcal{A}} + w_{\mathcal{B}}, \alpha_{\mathcal{A}} + \alpha_{\mathcal{B}}, \beta_{\mathcal{A}} + \beta_{\mathcal{B}}), \quad (4.63)$$

$$\mathcal{A} - \mathcal{B} = (w_{\mathcal{A}}, \alpha_{\mathcal{A}}, \beta_{\mathcal{A}}) - (w_{\mathcal{B}}, \alpha_{\mathcal{B}}, \beta_{\mathcal{B}}) = (w_{\mathcal{A}} - w_{\mathcal{B}}, \alpha_{\mathcal{A}} + \alpha_{\mathcal{B}}, \beta_{\mathcal{A}} + \beta_{\mathcal{B}}), \quad (4.64)$$

$$c \cdot \mathcal{A} = (cw_{\mathcal{A}}, c\alpha_{\mathcal{A}}, c\beta_{\mathcal{A}}), \quad (4.65)$$

przy czym  $\mathcal{A}$  i  $\mathcal{B}$  oznaczają liczby rozmyte, natomiast  $c$  – dodatnią liczbę rzeczywistą. Jeżeli przyjąć notację, w której liczba rzeczywista  $a$  jest zapisywana w postaci trzech parametrów  $a = (a, 0, 0)$ , to działania te mogą być uogólnione na dodawanie i odejmowanie liczby rozmytej i rzeczywistej. Co więcej, wzory (4.63)–(4.65) prawidłowo wyrażają wtedy także działania na dwóch liczbach rzeczywistych. Oznacza to, że liczba rozmyta typu L-R w powyższym zakresie jest uogólnieniem liczby rzeczywistej. Ostatecznie, każdy współczynnik  $w_i$ , wprowadzony we wzorze (4.20), może być uogólniony do dogodnej w identyfikacji i obliczeniach trójparametrycznej liczby rozmytej, oznaczanej poniżej jako  $\mathcal{W}_i = (w_i, \alpha_i, \beta_i)$ . W szczególnym przypadku  $\mathcal{W}_i = (w_i, 0, 0)$  reprezentuje liczbę ostrą  $w_i$ .



Jeżeli współczynniki  $w_i$  są rozmyte, to – jak wskazuje zależność (4.61) – wskaźnik jakości rozważanego układu stacji bazowych ma postać kombinacji liniowej trójparametrycznych liczb rozmytych  $\mathcal{W}_i$ , a zatem – z uwagi na wzory (4.63)–(4.65) – staje się on także trójparametryczną liczbą rozmytą, oznaczaną dalej jako  $\mathcal{E}$ . W celu umożliwienia porównywania jakości poszczególnych układów stacji bazowych, koniecznego przy wyznaczaniu rozwiązania optymalnego, użyta zostanie metodyka **teorii preferencji rozmytych** [5]. Funkcja preferencji  $\mathcal{P}$  liczby rozmytej  $\mathcal{E}$  o ograniczonym nośniku funkcji przynależności będzie przyjęta w postaci wynikłej z praktyki podejmowania decyzji (punkt 4.3.1):

$$\mathcal{P}(\mathcal{E}) = \delta \frac{\int_{\min \text{supp } \mu_{\mathcal{E}}}^{\max \text{supp } \mu_{\mathcal{E}}} x \mu_{\mathcal{E}}(x) dx}{\int_{\min \text{supp } \mu_{\mathcal{E}}}^{\max \text{supp } \mu_{\mathcal{E}}} \mu_{\mathcal{E}}(x) dx} + (1 - \delta) \min \text{supp } \mu_{\mathcal{E}}, \quad (4.66)$$

gdzie  $\delta \in [0, 1]$ ,  $\mu_{\mathcal{E}}$  oznacza funkcję przynależności liczby rozmytej  $\mathcal{E}$ , a  $\text{supp } \mu_{\mathcal{E}}$  – jej nośnik. Wartość funkcji przynależności jest zatem kombinacją liniową z wagami  $\delta$  oraz  $1 - \delta$  wartości średniej liczby rozmytej i wartości minimalnej jej nośnika. Wartość średnia odpowiada bayesowskiej regule decyzyjnej i wyraża działanie „realistyczne”, natomiast wartość minimalna nośnika funkcji przynależności wynika z reguły minimaksowej i reprezentuje działanie „pesymistyczne”. Parametr  $\delta$  stanowi zatem o strategii firmy operatorskiej w zakresie od realistycznej – zakładającej średni z przewidywanego popytu – dla  $\delta = 1$ , do pesymistycznej – zakładającej najmniejszy poziom przewidywanego popytu – dla  $\delta = 0$ . Jeśli brak jest wyraźnych preferencji, najczęściej wstępnie przyjmuje się  $\delta = 0,5$ .

Ostatecznie, jeżeli rozważane są dwa układy stacji bazowych scharakteryzowanych przez rozmyte wskaźniki jakości, to za „lepszy” należy uznać ten, dla którego wartość funkcji preferencji (4.66) jest większa.

Prezentowana metodyka wyznaczania optymalnego układu stacji bazowych systemu LMDS ma charakter uniwersalny i może być łatwo zaadaptowana do zagadnień pokrewnych, na przykład rozplanowania punktów obsługi klienta na terenie aglomeracji miejskich.

Niniejszy punkt ujmuje wyniki badań prowadzonych wraz z Jackiem Wa-głowskim, ze studiów doktoranckich Instytutu Badań Systemowych PAN, w ramach Jego pracy doktorskiej [31]. Powyższy tekst zawiera fragmenty tej pracy oraz wspólnych publikacji [25, 26].

## 4.4. Podsumowanie

W niniejszym rozdziale zostały przedstawione podstawy metodyki konstruowania estymatorów jądrowych i ich przykładowe zastosowania do zagadnień współczesnych badań systemowych. Ta podstawowa obecnie koncepcja estymacji nieparametrycznej, w swym pierwotnym ujęciu, pozwala oszacować gęstość rozkładu jedno- i wielowymiarowej zmiennej losowej, praktycznie niezależnie od jego postaci i cech. Dogodna formuła estymatora jądrowego umożliwia także uzyskanie, odmiennych od gęstości, funkcji charakteryzujących rozkład zmiennej

losowej, na przykład dystrybuanty. Co więcej, możliwe jest wyznaczanie parametrów rozkładu nierzadko z większą dokładnością niż w przypadku metod klasycznych. Koncepcja estymatorów jądrowych ma charakter uniwersalny i może być przeniesiona na wiele pokrewnych problemów, na przykład wyznaczania funkcji regresji lub gęstości widma.

Możliwości aplikacyjne estymatorów jądrowych są implikowane potencjałem obliczeniowym systemów komputerowych i dostępnością informacji, głównie wyrażanej licznnością próby losowej. Obecny postęp techniki numerycznej, jak również zautomatyzowanie procesów metrologicznych oraz pozyskiwania danych, powoduje ich ciągły i szybki wzrost. W niniejszym rozdziale możliwości aplikacyjne estymatorów jądrowych zostały zaprezentowane na czterech przykładowych zadaniach współczesnych badań systemowych, zilustrowanych problemami z zakresu archeologii, medycyny, sterowania automatycznego i telekomunikacji. Inne przykłady znaleźć można w pracach [10–17, 27].

Zaprezentowany powyżej materiał został szczegółowo przedstawiony w monografii [18]. Wiele aspektów można także znaleźć w klasycznych podręcznikach anglojęzycznych [30, 32].

## Literatura

- [1] Barnett V., Lewis T. (1994): *Outliers in Statistical Data*. Wiley, Chichester.
- [2] Berger J.O. (1980): *Statistical Decision Theory*. Springer, New York.
- [3] Charytanowicz M. (2005): Bayesowska determinizacja informacji nieprecyzyjnej w zagadnieniach medycznych. Praca doktorska, Instytut Badań Systemowych PAN, Warszawa.
- [4] Domański C., Pruska K. (2000): *Nieklasyczne metody statystyczne*. PWE, Warszawa.
- [5] Fodor J., Roubens M. (1994): *Fuzzy Preference Modelling and Multicriteria Decision Support*. Kluwer, Dordrecht.
- [6] Gajek L., Kałużka M. (2000): *Wnioskowanie statystyczne*. WNT, Warszawa.
- [7] Hryniewicz O. (2001): *Wykłady ze statystyki*. Wydawnictwo WSISiZ, Warszawa.
- [8] Kacprzyk J. (1986): *Zbiory rozmyte w analizie systemowej*. PWN, Warszawa.
- [9] Koronacki J., Mielniczuk J. (2001): *Statystyka*. WNT, Warszawa.
- [10] Kulczycki P. (1998): *Wykrywanie uszkodzeń w systemach zautomatyzowanych metodami statystycznymi*. Alfa, Warszawa.
- [11] Kulczycki P. (2000): A Statistical Fault Detection System. 16th IMACS World Congress on Scientific Computation, Applied Mathematics and Simulation, Lausanne (Szwajcaria), 21–25 sierpnia 2000, CD: 216–4.
- [12] Kulczycki P. (2000): Fuzzy Controller for Mechanical Systems. *IEEE Transactions on Fuzzy Systems*, Vol. 8, s. 645–652.
- [13] Kulczycki P. (2001): A Fuzzy Approach for Mechanical System with Uncertain Load. European Control Conference, Porto (Portugalia), 4–7 września 2001, s. 2658–2663, CD: ecc4561.
- [14] Kulczycki P. (2001): An Algorithm for Bayes Parameter Identification. *Journal of Dynamic Systems, Measurement, and Control*, Vol. 123, s. 611–614.
- [15] Kulczycki P. (2002): A test for comparing distributions functions with strongly unbalanced samples. *Statistica*, Vol. 62, s. 39–49.

- [16] Kulczycki P. (2002): Bayes Parameter Identification with Reference to Nonlinear Optimal Control. 15th World Congress IFAC, Barcelona (Hiszpania), 21–26 lipca 2002, CD: 1329.
- [17] Kulczycki P. (2002): Statistical Inference for Fault Detection: A Complete Algorithm Based on Kernel Estimators. *Kybernetika*, Vol. 38, s. 141–168.
- [18] Kulczycki P. (2005): *Estymatory jądrowe w analizie systemowej*. WNT, Warszawa.
- [19] Kulczycki P., Charytanowicz M. (2003): Zastosowanie bayesowskiej defuzyfikacji w zagadnieniach współczesnej medycyny. W: *Inżynieria Wiedzy i Systemy Ekspertowe*, Z. Bubnicki, A. Grzech (red.), Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław, Vol. 1, s. 91–98.
- [20] Kulczycki P., Charytanowicz M. (2005): Bayes sharpening of imprecise information. *Applied Mathematics and Computer Science*, Vol. 15, s. 393–404.
- [21] Kulczycki P., Mazgaj A. (2003): Parameter Identification with Nonsymmetrical Quadratic Loss Function for Optimal Control. 9th IEEE International Conference on Methods and Models in Automation and Robotics, Międzyzdroje, 25–28 sierpnia 2003, Vol. 1, s. 689–694, CD: I1-5.
- [22] Kulczycki P., Mazgaj A. (2005): An algorithm for Bayes parameter identification with quadratic asymmetrical loss function. *Control and Cybernetics*, Vol. 34, s. 1127–1148.
- [23] Kulczycki P., Prochot C. (2002): Identyfikacja stanów nietypowych za pomocą estymatorów jądrowych. W: *Metody i techniki analizy informacji i wspomaganie decyzji*, Z. Bubnicki, O. Hryniewicz, R. Kulikowski (red.), EXIT, Warszawa, s. V.57–62.
- [24] Kulczycki P., Prochot C. (2004): Wykrywanie elementów odosobnionych za pomocą metod estymacji nieparametrycznej. W: *Badania operacyjne i systemowe; podejmowanie decyzji – podstawy teoretyczne i zastosowania*, R. Kulikowski, J. Kacprzyk, R. Słowiński (red.), EXIT, Warszawa, s. 313–328.
- [25] Kulczycki P., Waglowski J. (2003): Optimal Base-Stations Locations in the LMDS Wireless Data Transmission System. First African Control Conference, Cape Town (RPA), 3–5 grudnia 2003, s. 375–380, CD: 108.
- [26] Kulczycki P., Waglowski J. (2005): On the application of statistical kernel estimators for the demand-based design of a wireless data transmission system. *Control and Cybernetics*, Vol. 34, s. 1149–1167.
- [27] Kulczycki P., Wisniewski R. (2002): Fuzzy Controller for a System with Uncertain Load. *Fuzzy Sets and Systems*, Vol. 131, s. 185–195.
- [28] Mazgaj A. (2005): Modelowanie niepewności parametrów obiektu dla potrzeb sterowania optymalnego. Praca doktorska, Akademia Górniczo-Hutnicza, Wydział Elektrotechniki, Automatyki, Informatyki i Elektroniki, Kraków.
- [29] Prochot C.: Identyfikacja stanów nietypowych w oparciu o metodykę estymatorów jądrowych. Praca doktorska, Instytut Badań Systemowych PAN, Warszawa (w przygotowaniu).
- [30] Silverman B.W. (1986): *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London.
- [31] Waglowski J. (2005): Metoda planowania optymalnego układu stacji bazowych bezprzewodowego systemu transmisji danych LMDS. Praca doktorska, Instytut Badań Systemowych PAN, Warszawa.
- [32] Wand M.P., Jones M.C. (1994): *Kernel Smoothing*. Chapman and Hall, London.