

Modelowanie i identyfikacja – laboratorium 6.

Nieparametryczna identyfikacja statycznych systemów nieliniowych.

Jądrowy estymator funkcji regresji

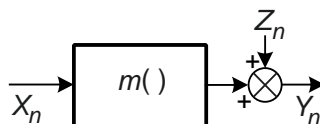
Paweł Wachel

Wymagania wstępne:

1. Wymagania wstępne z poprzednich zajęć¹.
2. Znajomość pojęć: system liniowy, nieliniowy, stacjonarny, niestacjonarny, dynamiczny, statyczny.
3. Znajomość konstrukcji i podstawowych własności jądrowego estymatora funkcji gęstości prawdopodobieństwa.

Zadania do wykonania:

1. Dany jest statyczny system nieliniowy z charakterystyką $m(x) = \text{atan}(a \cdot x)$, gdzie a jest pewną stałą (w początkowych eksperymentach przyjąć dla uproszczenia $a = 1$). Wygenerować N -elementowy sygnał wejściowy $\{X_n\}$ typu *i.i.d.* o rozkładzie $U[-2, 2]$



Rysunek 1: Statyczny system nieliniowy z addytywnym zakłóceniem na wyjściu

oraz niezależny od niego sygnał zakłócający $\{Z_n\}$ o rozkładzie normalnym $\mathcal{N}(0, \sigma_Z^2)$, σ_Z^2 jest dowolnie wybraną stałą.

2. Dysponując sekwencjami $\{X_n\}$ i $\{Z_n\}$ wyznaczyć odpowiadający im sygnał $\{Y_n\}$. Ciąg par obserwacji $\{(X_1, Y_1), (X_2, Y_2), \dots, (X_N, Y_N)\}$ będziemy oznaczać symbolem T_N . Wykreślić nieliniową charakterystykę systemu wraz z 'chmurą' pomiarów ze zbioru T_N (punkty na płaszczyźnie).
3. Zaimplementować jądrowy estymator funkcji regresji

$$\hat{m}_N(x) = \frac{\sum_{n=1}^N Y_n K\left(\frac{X_n - x}{h_N}\right)}{\sum_{n=1}^N K\left(\frac{X_n - x}{h_N}\right)}, \quad (1)$$

¹Całkujemy wiedzę... przynajmniej do wakacji.

gdzie $K(\cdot)$ jest funkcją jądra (por. wykład), a h_N jest parametrem wygładzania. Wykreślić estymator $\hat{m}_N(x)$ w funkcji x (na tle prawdziwej charakterystyki $m(x)$) dla ustalonej wartości N (np. $N = 500$), jądra prostokątnego i kilku przykładowych wartości parametru wygładzania h_N . Przedyskutować uzyskane wyniki.

4. Wykreślić estymator $\hat{m}_N(x)$ w funkcji x dla różnych funkcji jądra (np. dla jąder omawianych na wykładzie), ustalonej wartości N (np. $N = 500$) i ustalonego h_N . Przedyskutować uzyskane wyniki. Badania powtórzyć dla różnych wartości parametru a (patrz pkt. 1). Jaki wpływ na uzyskiwane wyniki estymacji ma kształt nieliniowości $m(\cdot)$ w systemie (jaki jest wpływ parametru a)?
5. Wyznaczyć wartość h , która minimalizuje błąd

$$valid(h) = \frac{1}{2Q} \sum_{q=-Q}^Q \left[\hat{m}_N\left(\frac{q}{Q}\right) - m\left(\frac{q}{Q}\right) \right]^2.$$

Przyjąć $Q = 100$.

6. Dla tak wybranego h wykreślić i zinterpretować wykresy z p. 3. i 4.
7. Przeprowadzone badania powtórzyć dla zakłócenia o rozkładzie Cauchy'ego $C(0, \gamma)$, gdzie $\gamma = 0.01$. Zmieniając parametr γ zbadać wpływ rozważanego zakłócenia na uzyskiwane rezultaty estymacji.

Zadania dodatkowe:

1. Dla wybranej gęstości rozkładu prawdopodobieństwa wejściowego $f(x)$ wygenerować L niezależnych, N -elementowych sekwencji pomiarowych typu wejście/wyjście i wyznaczyć błąd empiryczny

$$Err\{\hat{m}_N\} = \frac{1}{LM} \sum_{l=1}^L \sum_{m=1}^M \left[\hat{m}_N^{[l]}(x_m) - m(x_m) \right]^2,$$

w którym $\{x_1, x_2, \dots, x_M\}$ jest sekwencją równoodległych punktów z pewnego odcinka $[a, b]$. Przyjąć $M = 100$ oraz $L = 10$ i wykreślić $Err\{\hat{m}_N\}$ w funkcji h_N . Przedyskutować uzyskane wyniki.

2. Wykreślić błąd $Err\{\hat{m}_N\}$ w funkcji N . Następnie rozważyć zadanie identyfikacji, w którym $E\{Z_n\} = 1$ (niezerowa wartość oczekiwana błędu na wyjściu). Wykreślić błąd $Err\{\hat{m}_N\}$ i zinterpretować uzyskane wyniki.
3. Bazując na technice krosvalidacji omówionej w poprzednim ćwiczeniu zaproponować algorytm automatycznego strojenia estymatora \hat{m}_N , tj. algorytm doboru parametru wygładzania h_N .

Literatura:

1. Jakubowski Jacek, Sztencel Rafał. Wstęp do teorii prawdopodobieństwa. Script, 2001.
2. Wasserman, Larry. All of statistics: a concise course in statistical inference. Springer Science & Business Media, 2013.
3. Wasserman, Larry. All of nonparametric statistics. Springer Science & Business Media, 2006.
4. Plucińska Agnieszka, Pluciński Edmund. Probabilistyka: rachunek prawdopodobieństwa, statystyka matematyczna, procesy stochastyczne. Wydawnictwa Naukowo-Techniczne, 2000.
5. Gajek Lesław, Kałużka Marek. Wnioskowanie statystyczne: modele i metody. Wydawnictwa Naukowo-Techniczne, 1993.
6. Notatki z wykładu.