Ortogonalny estymator funkcji gęstości prawdopodobieństwa

dr hab. inż. Przemysław Śliwiński

23 kwietnia 2015

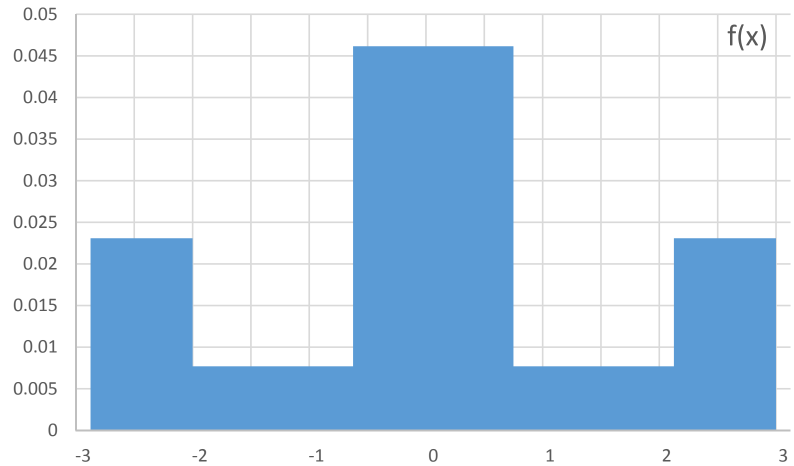
1 Estymacja ortogonalna

1. Wygenerować ciąg liczb $\{X_1, \ldots, X_N\}$, dla $N = 1024$ o wybranej gęstości[1]:
   (a) odcinkami stałej,
   (b) parzystej,
   (c) o nośniku w przedziale $[-\pi, \pi]$.

2. Zauważając, że
$$f(x) = \sum_{k=0}^{\infty} \beta_k \cos(kx),$$
gdzie
$$\beta_k = \langle f(\xi), \cos(k\xi) \rangle = \int_{-\pi}^{\pi} f(\xi) \cos(k\xi) \, d\xi = E[\cos(X)],$$
i gdzie $X \sim f(x)$, zaimplementować estymator ortogonalny tej gęstości o postaci
$$\hat{f}_K(x) = \sum_{k=0}^{K} \hat{\beta}_k \cos(kx), \text{ gdzie } \hat{\beta}_k = \frac{1}{N} \sum_{n=1}^{N} \cos(kX_n). \tag{1}$$

3. Wyznaczyć estymatory gęstości na podstawie podciągów dla $N = 128, 256, 512$ i 1024 dobierając parametr K (zwany parametrem odcięcia, ang. *cut-off parameter*) tak, aby uzyskać najmniejszy (empiryczny) błąd średniokwadratowy
$$emperror(h) = \sum_{q=-Q}^{Q} \left[f(x_q) - \hat{f}_K(x_q) \right]^2, \tag{2}$$
dla wybranego (z uzasadnieniem doboru!) ciągu $\{x_q\}, q = -Q, \ldots, Q)$.

[1] Np. metodą eliminacji.



Rysunek 1: Przykładowa gęstość

4. Jako alternatywę dla estymatora (1) rozważyć jego wersje z progowaniem (*threshold estimate*), w którym empiryczne współczynniki Fouriera $\hat{\beta}_k$ są progowane, tj. dla przyjętej wartości progu $t > 0$, wszystkie mniejsze od niego (co do modułu), są zerowane

$$\hat{f}_K^t(x) = \sum_{k=0}^{K} \hat{\beta}_k^t \cos(kx), \text{ gdzie } \hat{\beta}_k^t = \hat{\beta}_k \cdot \left(\left| \hat{\beta}_k \right| > t \right). \tag{3}$$

Zbadać zależność błędu (2) w zależności od K i t.