Estymatory ściągające dla funkcji masy prawdopodobieństwa

Małgorzata Łazęcka, Jan Mielniczuk

Wydział Matematyki i Nauk Informacyjnych PW
Instytut Podstaw Informatyki PAN

Rozważane będą sposoby estymacji funkcji masy prawdopodobieństwa p (w przypadku pary zmiennych dyskretnych (X, Y) ozn. $p = p(x, y)$). Przedstawiony zostanie estymator ściągający typu Jamesa-Steina postaci

$$\hat{p}_\lambda = \lambda \hat{p}^{(1)} + (1 - \lambda)\hat{p}^{(2)},$$

gdzie $p^{(1)}$ jest estymatorem związanym z modelem ograniczonym i jego wariancja jest niższa niż wariancja $p^{(2)}$ (np. jest etymatorem największej wiarogodności w modelu, w którym zakładamy niezależność zmiennych X i Y lub rozkład jednostajny na (X, Y)), $p^{(2)}$ jest estymatorem nieobciążonym, zaś parametr $\lambda \in [0, 1]$ jest dobierany w taki sposób, by uzyskany estymator p_λ minimalizował błąd średniokwadratowy.

W trakcie referatu zostanie przedstawiona propozycja nowego estymatora p_λ, w którym parametr λ_{SE} minimalizuje błąd kwadratowy:

$$\lambda_{SE} = \operatorname*{arg\,min}_{\lambda \in [0,1]} SE(\lambda) = \operatorname*{arg\,min}_{\lambda \in [0,1]} \sum_{x,y} \left(p(x, y) - \hat{p}_\lambda(x, y)\right)^2 = \frac{\sum_{x,y} \left(\hat{p}^{(2)}(x, y) - p(x, y)\right)\left(\hat{p}^{(2)}(x, y) - \hat{p}^{(1)}(x, y)\right)}{\sum_{x,y} \left(\hat{p}^{(1)}(x, y) - \hat{p}^{(2)}(x, y)\right)^2},$$

pokazany zostanie sposób estymacji λ_{SE} oraz zastosowanie estymatora $\hat{p}_{\hat{\lambda}_{SE}}$ do selekcji zmiennych opartej na miarach pochodzących z teorii informacji.

Literatura

[1] J. Hausser, K. Strimmer (2008), *Entropy Inference and the James-Stein Estimator, With Application to Nonlinear Gene Association Networks,* The Journal of Machine Learning Research 10, 1469–1484

[2] K. Sechidis, L. Azzimonti, A. Pocock, G. Corani, J. Weatherall, G. Brown (2019), *Efficient feature selection using shrinkage estimators,* Machine Learning 108, 1261–1286