6th sheet of exercises
due: 24 January 2008

Exercise 21. We consider the estimation of the first derivative of a density function f_X based on i.i.d data $X_1, \ldots, X_n$ having the density f_X. We restrict our consideration to the univariate case. Show that there is an l-order kernel (integer l) supported on $[-1,1]$, which is continuously differentiable on the whole real line, in addition. We choose the approach

$$\tilde{K}_l(x) \,=\, K(x) + \sum_{j=2l}^{2l+3} \alpha_j L_j(x)\,,\; x \in [-1,1]\,,$$

where K is the l-order kernel as constructed in the lectures and L_j denotes the Legendre polynomial with the degree j. Without any proof, one may use that L_j is an even or odd function if j is even or odd, respectively; $L_j(1) = 1$; and that $L_j'(1) < 0$ decreases monotonously when j increases.

Exercise 22. (Continuation of Exercise 21) As an estimator of $f_X'(x)$ we suggest to take the derivative of the kernel density estimator with the kernel function derived in Exercise 21. Give an upper bound on the convergence rates of the MSE in some $x \in \mathbb{R}$ uniformly over the Hölder classes $\mathcal{F}_{C,\beta;x}$ for $\beta > 1$.

Exercise 23. We consider the density class $\mathcal{F}_\omega$, $\omega > 0$, consisting of all univariate densities whose Fourier transforms are supported on $[-\omega,\omega]$. Apply the kernel density estimator with the sinc kernel $K(x) = (\sin x)/(\pi x)$ to estimate $f_X \in \mathcal{F}_\omega$ based on i.i.d. data $X_1, \ldots, X_n$ with the density f_X. Show that the parametric rates n^{-1} are achievable with respect to the MISE, considered uniformly over the density class $\mathcal{F}_\omega$, by appropriate bandwidth selection. Why is this choice of the bandwidth no contradiction to Devroye's equivalence theorem (Theorem 3.1)?

Exercise 24. We define the smoothness class

$$\mathcal{G}_{C_1,C_2,\gamma} \,=\, \Big\{ f \text{ univariate density} \,:\, \int \big|f^{ft}(t)\big|^2 \exp(C_1|t|^\gamma)\, dt \,\leq\, C_2 \Big\}\,,$$

with $C_1, C_2, \gamma > 0$. Determine the convergence rates of

$$\sup_{f_X \in \mathcal{G}_{C_1,C_2,\gamma}} E\|\hat{f}_X - f_X\|_2^2\,,$$

for the kernel density estimator $\hat{f}_X$ under optimal selection of the kernel function K and the bandwidth h.