Statistical Theory MATH-442
Spring 2025 Ho Yun & Yoav Zemel

Exercise sheet 11

Exercise 1 Here we construct a kernel K of arbitrary order $\ell+1$ such that K is bounded, C^∞, and supported on $[-1,1]$.

Consider the bump function $g(z) = e^{-1/(1-z^2)}$ on $[-1,1]$ and 0 otherwise, which is supported on $[-1,1]$, strictly positive on $(-1,1)$, and C^∞ on all of $\mathbb{R}$.

Define $K(z) = \sum_{j=0}^{M} a_j z^j g(z)$. Show that it is possible to choose M and $(a_j)_{j=0}^{M}$ such that K satisfies the desired properties. *Hint: the constraints define $\ell+1$ linear equations on the a_j's. Write them as a matrix and show that for $M=\ell$ this matrix is strictly positive definite, thus invertible.*

Exercise 2 Let $f \in C_{den}^{\beta}(M)$, and let K be any bounded kernel of order at least β (the previous exercise shows that such K exists). Show that $\widehat{f}_h(x)$ is bounded by a constant depending only on h and K. Using the bound on the bias with $h=1$, show that $f(x)$ is bounded by a constant depending only on K, β, and M. What is the best dependence on M you can get?

Exercise 3 The following result is sometimes shown in textbooks. Let f be a twice differentiable density function. If K is a kernel of order 2 with $R(K) < \infty$ and $\mu_2(|K|) < \infty$, then

$$MSE\left(\widehat{f}_h(x)\right) = \frac{R(K)f(x)}{nh} + \frac{1}{4}h^4\mu_2^2(K)f''(x)^2 + o\left(\frac{1}{nh} + h^4\right).$$

(If you like an easier problem, assume that the second derivative is bounded. Otherwise, use directly the definition that $f(x+y) - f(x) - yf'(x) - y^2 f''(x)/2 = o(y^2)$ as $y \to 0$.)

Exercise 4 The previous exercise suggests that we should choose $h = n^{-1/5}$ to obtain a mean squared error of the order $n^{-4/5}$, consistent with the results we had for general β ($\beta = 2$ here). We are still left to choose the kernel K. Since the pair (K,h) is equivalent to the pair $(K_h, 1)$, we need to fix the scale of K in some manner. We shall do this by choosing $\mu_2(|K|) = 1$. Furthermore, we shall restrict attention to nonnegative kernels K. Then, the bound on the mean squared error from the previous exercise suggests we need to minimise $R(K) = \int K^2(z)dz$ subject to $\int K(z)dz = 1 = \int K(z)z^2 dz$. Let us see how this optimal kernel K looks like. It is called the *Epanechnikov* kernel.

Remark. The most important part in this exercise is the last one, and it can be done independently of the rest of the question. The rest of the exercise tries to explain where the result comes from, rather than give it out of the blue.

1. Let V be any function such that $\int V(z)dz = 0 = \int z^2 V(z)dz$. Show informally that $\int K(z)V(z)dz = 0$. *Hint:* argue informally that $R(K) \le R(K+tV)$ for all $t \in \mathbb{R}$. This is informal, since in principle we need to also impose that $K + Vt$ is a nonnegative function; pretend that this restriction is not necessary.

2. Since V is ortohgonal to the functions 1 and z^2, argue informally that K must be in the span of these functions.

3. Since a function of the form $a + bz^2$ will never be integrable (unless $a = 0 = b$), we truncate it to be symmetric and with compact support $[-c,c]$ for some $c > 0$. Thus, our candidate kernel is $K(z) = (a+bz^2)1(|z| \le c)$. Find a and b as a function of c, and find conditions on c such that K is nonnegative.

4. One can write the objective function $R(K)$ as a function only of c, and show that it is nondecreasing (you do **not** need to do this, but it is at least very easy to show that $R(K) = O(1/c)$ as $c \to \infty$). Conclude that

$$K(z) = \tfrac{3}{4\sqrt{5}} \left(1 - \tfrac{z^2}{5}\right) 1(|z| \leq 5).$$

5. Now that we have figured out the answer informally, prove rigorously that K from the previous part is indeed the minimiser. *Hint:* Compare K with $K+V$, where this time assume that $K + V$ is nonnegative on $\mathbb{R}$.