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.*

Solution 1 Clearly K is supported on $[-1,1]$ and is C^∞ there. Define

$$m_j=\int_{-1}^{1} z^j g(z)dz.$$

($m_j=0$ for j odd but we will not use this property.) Then the conditions become

$$\int K(z)z^i dz=\sum_{j=0}^{M} a_j m_{i+j}=\begin{cases}1 & i=0\\ 0 & i\in\{1,\dots,\ell\}.\end{cases}$$

If we define a matrix $B\in\mathbb{R}^{(\ell+1)\times(M+1)}$ with coordinates $(B_{ij})=m_{i+j}$, then the conditions are that $B(a_0,\dots,a_M)^\top=(1,0,\dots,0)^\top\in\mathbb{R}^{\ell+1}$. So all we need to show that B can be made injective. Obviously this cannot happen unless $M\geq\ell$, so we shall try with $M=\ell$. Then B is a symmetric square matrix and will in fact be positive definite since

$$\sum_{i,j=0}^{\ell} a_i B_{ij} a_j=\sum_{i,j=0}^{\ell} a_i a_j m_{i+j}=\sum_{i,j=0}^{\ell} a_i a_j\int_{-1}^{1} z^{i+j} g(z)dz=\int_{-1}^{1}(\sum_{j=0}^{\ell} a_j z^j)^2 g(z)dz.$$

Since g is strictly positive, this is integral is positive unless $a_j=0$ for all j. Thus B is positive definite and invertible, so there exists a choice $(a_j)_{j=0}^{\ell}$ that makes K a kernel with all the desired properties.

Exercise 2 Let $f \in C^{\beta}_{den}(M)$, 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?

Solution 2 Obviously

$$|\widehat{f}_h(x)| = \left| \tfrac{1}{nh} \sum_{i=1}^{n} K\left(\tfrac{x - X_i}{h}\right) \right| \le \tfrac{1}{h} \|K\|_\infty.$$

Therefore $|\mathbb{E}\widehat{f}_h(x)|$ is bounded by the same value. For $f \in C^{\beta}_{den}(M)$, we have seen that

$$|\mathbb{E}\widehat{f}_1(x) - f(x)| \le \tfrac{M}{\ell!} \int |z|^\beta |K(z)| dz, \qquad \ell = \lceil \beta \rceil - 1.$$

It follows that for all $x \in \mathbb{R}$

$$f(x) \le |\mathbb{E}\widehat{f}_1(x)| + |\mathbb{E}\widehat{f}_1(x) - f(x)| \le \|K\|_\infty + \tfrac{M}{\ell!} \int |z|^\beta |K(z)| dz$$

which depends only on K, β and M.

Instead of $h = 1$ we can optimise over h the bound

$$f(x) \le |\mathbb{E}\widehat{f}_h(x)| + |\mathbb{E}\widehat{f}_h(x) - f(x)| \le \tfrac{1}{h} \|K\|_\infty + h^\beta \tfrac{M}{\ell!} \int |z|^\beta |K(z)| dz$$

by choosing $h = M^{-1/(\beta+1)}$ and then the obtained upper bound behaves like $M^{1/(\beta+1)}$, which is consistent with the bound shown in class for $\beta = 2$.

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

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

(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 $y \to 0$.)

Solution 3 As in the proof in lectures define

$$R_1(x,z,h) = f(x-hz) - f(x) + hzf'(x) - \tfrac{1}{2}h^2z^2f''(x),$$

and since $\mu_1(K) = 0$ we have

$$\mathbb{E}\widehat{f}_h(x) - f(x) = \tfrac{h^2\mu_2(k)f''(x)}{2} + \int K(z)R_1(x,z,h)dz.$$

Differentiability means that for any ϵ there exists δ such that if $|hz| < \delta$ then $|R_1(x,h,z)| \le \epsilon h^2 z^2$. Thus

$$|\int K(z)R_1(x,z,h)dz| \le \int_{|z|\le\delta/h} \epsilon h^2|K(z)|z^2dz + \int_{|z|>\delta/h} |K(z)||f(x-hz) - f(x) + hzf'(x) - \tfrac{1}{2}h^2z^2f''(x)|dz.$$

The first term is bounded by $\epsilon h^2\mu_2(|K|)$. To bound the second integral, we multiply by $z^2h^2/\delta^2 \ge 1$ or by $|z|h/\delta \ge 1$ so as to always get $K(z)z^2$ in the integrals:

$$\begin{aligned}\int_{|z|>\delta/h} |K(z)||f(x-hz) - f(x)|dz &\le 2\|f\|_\infty \tfrac{h^2}{\delta^2}\int_{|z|>\delta/h} |K(z)|z^2dz\\ \int_{|z|>\delta/h} |K(z)||hzf'(x)| &\le |f'(x)|\tfrac{h^2}{\delta}\int_{|z|>\delta/h} |K(z)|z^2dz\\ \int_{|z|>\delta/h} |K(z)||\tfrac{1}{2}h^2z^2f''(x)| &\le \tfrac{|f''(x)|}{2}h^2\int_{|z|>\delta/h} |K(z)|z^2dz\end{aligned}$$

Since δ is fixed and $h \to 0$, the integral at the right hand side of the three lines vanishes (by dominated convergence and $\mu_2(|K|) < \infty$). Therefore the bias is of order $h^2\mu_2(K)f''(x)/2 + o(h^2)$.

As for the variance, we have

$$\text{Var}\widehat{f}_h(x) = \tfrac{1}{nh^2}\int f(y)K^2((x-y)/h)dy - \tfrac{1}{n}(\mathbb{E}\widehat{f}_h(x))^2,$$

and we already know that expectation is $f(x) + O(h^2)$ so the negative contribution is $O(1/n)$. For the integral we apply the usual change of variables $hz = x - y$ to obtain

$$\tfrac{1}{nh}\int f(y)K^2((x-y)/h)dy = \tfrac{1}{nh}R(K)f(x) + \tfrac{1}{nh}\int(f(x-zh) - f(x))K^2(z)dz.$$

The last integral is $o(1)$ by a similar (simpler) truncation argument, since f is continuous at x. Thus:

$$\text{Var}\widehat{f}_h(x) = \tfrac{R(k)f(x)}{nh} + O(\tfrac{1}{n}) + o(\tfrac{1}{nh}) = \tfrac{R(k)f(x)}{nh} + o(\tfrac{1}{nh})$$

Thus we have for the mean squared error

$$MSE(\widehat{f}_h(x)) = \text{Var}\widehat{f}_h(x) + (\mathbb{E}\widehat{f}_h(x) - f(x))^2 = \tfrac{R(k)f(x)}{nh} + \tfrac{1}{n}h^4\mu_2(k)^2(f''(x))^2 + o(\tfrac{1}{nh}) + o(h^4).$$

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 $\int K(z)dz = 1 = \int K(z)z^2dz$. 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 $\int V(z)dz = 0 = \int z^2V(z)dz$. Show informally that $\int K(z)V(z)dz = 0$. *Hint:* argue informally that $R(K) \leq R(K + tV)$ 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 $K(z) = (a + bz^2)1(|z| \leq c)$. 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 $\mathbb{R}$.

Solution 4

1. Ignoring the positivity issue, since $K + tV$ also satisfies the constraints and K is optimal, we have
$$R(K) \leq R(K + tV) = \int K^2(z)dz + 2t\int K(z)V(z)dz + t^2\int V^2(z)dz$$
for all $t \in \mathbb{R}$. This is only possible if $\int K(z)V(z) = 0$.

2. K is orthogonal to all such V, so it is orthogonal to $(\mathit{span}\{1, z^2\})^\perp$. Therefore $K \in \mathit{span}\{1, z^2\}$.

3. We have the two equations
$$1 = \int K(z)dz = 2ac + \tfrac{2bc^3}{3}$$
$$1 = \int K(z)z^2dz = \tfrac{2ac^3}{3} + \tfrac{2bc^5}{5}$$

from which we obtain

$$a = \tfrac{9c^2-15}{8c^3} \qquad b = \tfrac{3-6ac}{2c^3}.$$

Now K is nonnegative if and only if $a \geq 0$ and $a + bc^2 \geq 0$, which happen if and only if $\sqrt{15/9} \leq c \leq \sqrt{5}$.

4. Since $R(K)$ is decreasing, we take the maximal possible c, which is $c = \sqrt{5}$. This gives $a = 30/40\sqrt{5}$ and $a + 5b = 0$.

5. We have

$$R(K+V) = R(K) + R(V) + 2\int K(z)V(z)dz,$$

and it suffices to show that the last integral is nonnegative. By the formula for K and the conditions on V we have

$$\tfrac{4\sqrt{5}}{3}\int K(z)V(z)dz = \int_{-\sqrt{5}}^{\sqrt{5}} V(z)(1-z^2/5)dz = \int_{|z|\geq\sqrt{5}} \left(\tfrac{z^2}{5}-1\right)V(z)dz \geq 0,$$

where the inequality follows from $0 \leq K(z) + V(z) = V(z)$ for $|z| \geq \sqrt{5}$ and the equality before that from $\int V(z)dz = 0 = \int V(z)z^2dz$.