Solutions to Assignment 4

Solutions to problems 1,2 and 3 were obtained from Qirong Ho.

1

(a)

We seek a bound on the bias of $\hat{p}_{n,s}(x)$ at a point x_0:

$$\begin{aligned}
\mathbb{E}\left[\hat{p}_{n,s}(x_0) - p^{(s)}(x_0)\right] &\le \left|\mathbb{E}\left[\hat{p}_{n,s}(x_0) - p^{(s)}(x_0)\right]\right| \\
&= \left|\mathbb{E}\left[\frac{1}{nh^{s+1}}\sum_{i=1}^{n} K\left(\frac{X_i - x_0}{h}\right) - p^{(s)}(x_0)\right]\right| \\
&= \left|\frac{1}{h^{s+1}}\mathbb{E}\left[K\left(\frac{X - x_0}{h}\right)\right] - p^{(s)}(x_0)\right| \\
&= \left|\frac{1}{h^{s+1}}\int K\left(\frac{x - x_0}{h}\right) p(x)\, dx - p^{(s)}(x_0)\right|
\end{aligned}$$

Let $t = \frac{x-x_0}{h}$, $\frac{dt}{dx} = \frac{1}{h}$:

$$\begin{aligned}
&= \left|\frac{1}{h^s}\int K(t)\, p(th + x_0)\, dt - p^{(s)}(x_0)\right| \\
&= \left|\int K(t)\, h^{-s}\left(p(th + x_0) - p_{x_0,\beta}(th + x_0)\right) dt + \int K(t)\, h^{-s} p_{x_0,\beta}(th + x_0)\, dt - p^{(s)}(x_0)\right| \\
&\le \left|\int K(t)\, h^{-s}\left(p(th + x_0) - p_{x_0,\beta}(th + x_0)\right) dt\right| + \left|\int K(t)\, h^{-s} p_{x_0,\beta}(th + x_0)\, dt - p^{(s)}(x_0)\right| \qquad (1)
\end{aligned}$$

where $p_{x_0,\beta}(u)$ is the $\lfloor\beta\rfloor$-order Taylor expansion of $p(u)$ at a point x_0:

$$p_{x_0,\beta}(u) = \sum_{|i|\le\lfloor\beta\rfloor} \frac{(u - x_0)^i}{i!} p^{(i)}(x_0)$$

Since $p \in \Sigma(\beta, L)$, we have that

$$\left|p(th + x_0) - p_{x_0,\beta}(th + x_0)\right| \le L\,|th|^{\beta}$$

and therefore the first term in Eq. (1) is bounded:

$$\begin{aligned}
\left|\int K(t)\, h^{-s}\left(p(th + x_0) - p_{x_0,\beta}(th + x_0)\right) dt\right| &\le \left|\int K(t)\, h^{-s} L\,|th|^{\beta}\, dt\right| \\
&= h^{\beta - s} L \int K(t)\,|t|^{\beta}\, dt \\
&= h^{\beta - s} c
\end{aligned}$$

where $c = L \int K(t) |t|^{\beta} \, dt > 0$. Next, observe that $p_{x_0,\beta}(th + x_0)$ is polynomial in t with degree $\lfloor \beta \rfloor$. In particular, the s-th order term is $\frac{(th)^s}{s!} p^{(s)}(x_0)$. Hence

$$\begin{aligned}
\int K(t) h^{-s} p_{x_0,\beta}(th + x_0) \, dt &= \int K(t) h^{-s} \frac{(th)^s}{s!} p^{(s)}(x_0) \, dt \quad \text{(only the } s\text{-th moment of } K \text{ is nonzero)} \\
&= p^{(s)}(x_0) \int K(t) \frac{t^s}{s!} \, dt \\
&= p^{(s)}(x_0)
\end{aligned}$$

and therefore the second term in Eq. (1) is zero. We therefore conclude that

$$\mathbb{E}\left[\hat{p}_{n,s}(x_0) - p^{(s)}(x_0)\right] \leq h^{\beta-s} c \quad \forall p \in \Sigma(\beta, L)$$

and thus

$$\sup_{p \in \Sigma(\beta, L)} \mathbb{E}\left[\hat{p}_{n,s}(x_0) - p^{(s)}(x_0)\right] \leq h^{\beta-s} c$$

(b)

We seek a bound on the variance of $\hat{p}_{n,s}(x)$ at a point x_0. Let

$$Z_i = \frac{1}{h^{s+1}} K\left(\frac{X_i - x_0}{h}\right)$$

so that

$$\hat{p}_{n,s}(x_0) = \frac{1}{n} \sum_{i=1}^{n} Z_i$$

Then

$$\begin{aligned}
\mathbb{V}(Z_i) = \mathbb{V}(Z) &\leq \mathbb{E}\left(Z^2\right) \\
&= \frac{1}{h^{2s+2}} \int K^2\left(\frac{x - x_0}{h}\right) p(x) \, dx
\end{aligned}$$

Let $t = \frac{x - x_0}{h}$, $\frac{dt}{dx} = \frac{1}{h}$:

$$= \frac{1}{h^{2s+1}} \int K^2(t) \, p(th + x_0) \, dt$$

$p(x) \in \Sigma(\beta, L)$ implies that $p(x)$ has a bounded first derivative (assuming $\beta \geq 1$), and therefore $\sup_{x_0} p(x_0)$ is finite. Hence we can bound $p(th + x_0)$ inside the integral by $\sup_{x_0} p(x_0)$:

$$\begin{aligned}
&\leq \frac{1}{h^{2s+1}} \int K^2(t) \sup_{x_0} p(x_0) \, dt \\
&= \frac{\sup_{x_0} p(x_0)}{h^{2s+1}} \int K^2(t) \, dt \\
&= \frac{c'}{h^{2s+1}}
\end{aligned}$$

where $c' = \sup_{x_0} p(x_0) \int K^2(t) \, dt > 0$. Thus

$$\begin{aligned}
\mathbb{V}(\hat{p}_{n,s}(x_0)) &= \mathbb{V}\left(\frac{1}{n} \sum_{i=1}^{n} Z_i\right) \\
&\leq \frac{c'}{nh^{2s+1}} \quad \forall p \in \Sigma(\beta, L)
\end{aligned}$$

and therefore

$$\sup_{p \in \Sigma(\beta, L)} \mathbb{V}(\hat{p}_{n,s}(x_0)) \leq \frac{c'}{nh^{2s+1}}$$

(c)

Combining (a), (b) gives us a bound on the MSE:

$$\begin{aligned}
\sup_{p\in\Sigma(\beta,L)} \mathbb{E}\left[\left(\hat{p}_{n,s}(x_0) - p^{(s)}(x_0)\right)^2\right] &= \sup_{p\in\Sigma(\beta,L)} \text{Bias}^2\left(\hat{p}_{n,s}(x_0)\right) + \mathbb{V}\left(\hat{p}_{n,s}(x_0)\right) \\
&\leq h^{2\beta-2s}c^2 + \frac{c'}{nh^{2s+1}} \quad \forall x_0 \in \mathbb{R}
\end{aligned}$$

and therefore

$$\sup_{x_0\in\mathbb{R}} \sup_{p\in\Sigma(\beta,L)} \mathbb{E}\left[\left(\hat{p}_{n,s}(x_0) - p^{(s)}(x_0)\right)^2\right] \leq h^{2\beta-2s}c^2 + \frac{c'}{nh^{2s+1}}$$

Choosing the optimal bandwidth implies minimizing the RHS wrt h:

$$\begin{aligned}
\frac{d}{dh}h^{2\beta-2s}c^2 + \frac{c'}{nh^{2s+1}} &= 0 \\
(2\beta - 2s)\, h^{2\beta-2s-1}c^2 + \frac{(-2s-1)\,c'}{nh^{2s+2}} &= 0 \\
(2\beta - 2s)\, h^{2\beta-2s-1}c^2 &= \frac{(2s+1)\,c'}{nh^{2s+2}} \\
h^{2\beta+1} &= \frac{1}{n} \cdot \frac{(2s+1)\,c'}{c^2\,(2\beta-2s)} \\
h &= n^{-\frac{1}{2\beta+1}} C_0
\end{aligned}$$

where $C_0 = \left(\frac{(2s+1)c'}{c^2(2\beta-2s)}\right)^{1/(2\beta+1)}$. Hence

$$\begin{aligned}
\sup_{x_0\in\mathbb{R}} \sup_{p\in\Sigma(\beta,L)} \mathbb{E}\left[\left(\hat{p}_{n,s}(x_0) - p^{(s)}(x_0)\right)^2\right] &\leq \left(n^{-\frac{1}{2\beta+1}} C_0\right)^{2\beta-2s} c^2 + \frac{c'}{n\left(n^{-\frac{1}{2\beta+1}} C_0\right)^{2s+1}} \\
&= n^{-\frac{2(\beta-s)}{2\beta+1}} \left(C_0^{2(\beta-s)} c^2\right) + \frac{c'}{nn^{-\frac{2s+1}{2\beta+1}} C_0^{2s+1}} \\
&= n^{-\frac{2(\beta-s)}{2\beta+1}} \left(C_0^{2(\beta-s)} c^2\right) + \frac{c'}{n^{-\frac{2s-2\beta}{2\beta+1}} C_0^{2s+1}} \\
&= n^{-\frac{2(\beta-s)}{2\beta+1}} \left(C_0^{2(\beta-s)} c^2 + \frac{c'}{C_0^{2s+1}}\right) \\
&= Cn^{-\frac{2(\beta-s)}{2\beta+1}}
\end{aligned}$$

where $C = C_0^{2(\beta-s)}c^2 + \frac{c'}{C_0^{2s+1}}$. The above argument holds for all n, so it also holds as $n \to \infty$.

(d)

We are given

$$\begin{aligned}
K(u) &= \sum_{m=0}^{\lfloor\beta\rfloor} \phi_m^{(s)}(0)\, \phi_m(u)\, \mathbb{I}\{|u| \leq 1\} \\
\phi_0(x) &= \frac{1}{\sqrt{2}}, \quad \forall x \in [-1,1] \\
\phi_m(x) &= \sqrt{\frac{2m+1}{2}} \frac{1}{2^m m!} \frac{d^m}{dx^m}\left[\left(x^2-1\right)^m\right], \quad \forall m \geq 1, x \in [-1,1] \\
\int_{-1}^{1} \phi_j(x)\, \phi_k(x)\, dx &= \delta_{jk}
\end{aligned}$$

Since $\{\phi_m\}_{m=0}^{\infty}$ is an orthonormal basis, we can write

$$\begin{aligned} u^j &= \sum_{i=0}^{\infty} \theta_i \phi_i(u) \\ \theta_i &= \int_{-1}^{1} u^j \phi_i(u)\, du \end{aligned}$$

Hence

$$\begin{aligned} \int u^j K(u)\, du &= \int_{-1}^{1} u^j \sum_{m=0}^{\lfloor\beta\rfloor} \phi_m^{(s)}(0)\, \phi_m(u)\, du \\ &= \sum_{m=0}^{\lfloor\beta\rfloor} \int_{-1}^{1} u^j \phi_m^{(s)}(0)\, \phi_m(u)\, du \\ &= \sum_{m=0}^{\lfloor\beta\rfloor} \int_{-1}^{1} \left(\sum_{i=0}^{\infty} \theta_{ji} \phi_i(u) \right) \phi_m^{(s)}(0)\, \phi_m(u)\, du \\ &= \sum_{m=0}^{\lfloor\beta\rfloor} \sum_{i=0}^{\infty} \theta_{ji} \phi_m^{(s)}(0) \int_{-1}^{1} \phi_i(u)\, \phi_m(u)\, du \\ &= \sum_{m=0}^{\lfloor\beta\rfloor} \theta_m \phi_m^{(s)}(0) \end{aligned}$$

We now consider cases:

- Case 1 $j < s$. Observe that $\phi_m^{(s)}(0) = 0$ whenever $m < s$, since ϕ_m is an m-th degree polynomial. Hence

  $$\int u^j K(u)\, du = \sum_{m=s}^{\lfloor\beta\rfloor} \theta_m \phi_m^{(s)}(0)$$

  Because u^j is a j-th degree polynomial, $\theta_i = 0$ for all $i > j$. Therefore

  $$\int u^j K(u)\, du = \sum_{m=s}^{j} \theta_m \phi_m^{(s)}(0)$$

  But $j < s$, so the sum is empty, and therefore equals zero.

- Case 2 $j = s$. Observe that

  $$\begin{aligned} u^j &= \sum_{i=0}^{j} \theta_i \phi_i(u) \\ \frac{d^j}{du^j} u^j &= \frac{d^j}{du^j} \sum_{i=0}^{j} \theta_i \phi_i(u) \\ j! &= \sum_{i=0}^{j} \theta_i \phi_i^{(j)}(u) \\ j! &= \sum_{i=0}^{j} \theta_i \phi_i^{(j)}(0) \quad (\phi_i^{(j)}(u) \text{is constant for all } i \le j) \end{aligned}$$

and therefore when $j = s$,

$$\begin{aligned}
\int u^j K(u)\, du &= \int_{-1}^{1} \left(\sum_{i=0}^{s} \theta_i \phi_i(u) \right) \sum_{m=0}^{\lfloor \beta \rfloor} \phi_m^{(s)}(0)\, \phi_m(u)\, du \\
&= \sum_{i=0}^{s} \theta_i \phi_i^{(s)}(0) \\
&= s!
\end{aligned}$$

- Case 3 $j > s$. Observe that

$$\begin{aligned}
u^j &= \sum_{i=0}^{j} \theta_i \phi_i(u) \\
\frac{d^s}{du^s} u^j &= \frac{d^s}{du^s} \sum_{i=0}^{j} \theta_i \phi_i(u) \\
\frac{j!}{(j-s)!} u^{j-s} &= \sum_{i=0}^{j} \theta_i \phi_i^{(s)}(u)
\end{aligned}$$

When $u = 0$, we get

$$0 = \sum_{i=1}^{j} \theta_i \phi_i^{(s)}(0)$$

and hence

$$\begin{aligned}
\int u^j K(u)\, du &= \int_{-1}^{1} \left(\sum_{i=0}^{j} \theta_i \phi_i(u) \right) \sum_{m=0}^{\lfloor \beta \rfloor} \phi_m^{(s)}(0)\, \phi_m(u)\, du \\
&= \sum_{i=0}^{j} \theta_i \phi_i^{(s)}(0) \\
&= 0
\end{aligned}$$

2

(a)

We handle the proof by cases:

- Clearly the equation equals 1 when both $j, k = 1$.
- Suppose WLOG that $j = 1$, k is even. Then

$$\frac{1}{n} \sum_{s=1}^{n} \varphi_j(s/n)\, \varphi_k(s/n) = \frac{\sqrt{2}}{n} \sum_{s=1}^{n} \cos\left(\frac{\pi k s}{n} \right) = 0$$

- Suppose WLOG that $j = 1$, $k > 1$ is odd. Then

$$\frac{1}{n} \sum_{s=1}^{n} \varphi_j(s/n)\, \varphi_k(s/n) = \frac{\sqrt{2}}{n} \sum_{s=1}^{n} \sin\left(\frac{\pi (k-1) s}{n} \right) = 0$$

- Suppose that j, k are even. If $j \neq k$, then

$$\begin{aligned}\frac{1}{n}\sum_{s=1}^{n}\varphi_j(s/n)\,\varphi_k(s/n) &= \frac{1}{n}\sum_{s=1}^{n}2\cos\left(\frac{\pi js}{n}\right)\cos\left(\frac{\pi ks}{n}\right)\\ &= \frac{1}{n}\sum_{s=1}^{n}\cos\left(\frac{\pi(j-k)s}{n}\right)+\frac{1}{n}\sum_{s=1}^{n}\cos\left(\frac{\pi(j+k)s}{n}\right)\\ &= 0+0=0\end{aligned}$$

  If $j = k$, then

$$\begin{aligned}\frac{1}{n}\sum_{s=1}^{n}\varphi_j(s/n)\,\varphi_k(s/n) &= \frac{1}{n}\sum_{s=1}^{n}\cos(0)+\frac{1}{n}\sum_{s=1}^{n}\cos\left(\frac{2\pi js}{n}\right)\\ &= 1+0=1\end{aligned}$$

- Suppose that $j, k > 1$ are odd. Then

$$\begin{aligned}\frac{1}{n}\sum_{s=1}^{n}\varphi_j(s/n)\,\varphi_k(s/n) &= \frac{1}{n}\sum_{s=1}^{n}2\sin\left(\frac{\pi(j-1)s}{n}\right)\sin\left(\frac{\pi(k-1)s}{n}\right)\\ &= \frac{1}{n}\sum_{s=1}^{n}\cos\left(\frac{\pi(j-k)s}{n}\right)-\frac{1}{n}\sum_{s=1}^{n}\cos\left(\frac{\pi(j+k-2)s}{n}\right)\\ &= 0-0=0\end{aligned}$$

  If $j = k$, then

$$\begin{aligned}\frac{1}{n}\sum_{s=1}^{n}\varphi_j(s/n)\,\varphi_k(s/n) &= \frac{1}{n}\sum_{s=1}^{n}\cos(0)-\frac{1}{n}\sum_{s=1}^{n}\cos\left(\frac{2\pi(j-1)s}{n}\right)\\ &= 1-0=1\end{aligned}$$

- Finally, suppose WLOG that j is even and $k > 1$ is odd. Then

$$\begin{aligned}\frac{1}{n}\sum_{s=1}^{n}\varphi_j(s/n)\,\varphi_k(s/n) &= \frac{1}{n}\sum_{s=1}^{n}2\cos\left(\frac{\pi js}{n}\right)\sin\left(\frac{\pi(k-1)s}{n}\right)\\ &= \frac{1}{n}\sum_{s=1}^{n}\sin\left(\frac{\pi(j+k-1)s}{n}\right)-\frac{1}{n}\sum_{s=1}^{n}\sin\left(\frac{\pi(j-k+1)s}{n}\right)\\ &= 0-0=0\end{aligned}$$

(b)

$$\begin{aligned}\mathbb{E}\left[\hat{\theta}_j-\theta_j\right] &= \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n}Y_i\varphi_j(X_i)\right]-\int_0^1 f(x)\,\varphi_j(x)\,dx\\ &= \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n}\left(f(X_i)+\epsilon_i\right)\varphi_j(X_i)\right]-\int_0^1 f(x)\,\varphi_j(x)\,dx\\ &= \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n}f(X_i)\,\varphi_j(X_i)\right]-\int_0^1 f(x)\,\varphi_j(x)\,dx \quad (\mathbb{E}[\epsilon_i]=0)\\ &= \frac{1}{n}\sum_{i=1}^{n}f(i/n)\,\varphi_j(i/n)-\int_0^1 f(x)\,\varphi_j(x)\,dx \quad (X_i=i/n\ \text{ are nonrandom})\\ &= \alpha_j\end{aligned}$$

(c)

$$\begin{aligned}
\mathbb{E}\left[\left(\hat{\theta}_j - \mathbb{E}\left[\hat{\theta}_j\right]\right)^2\right] &= \mathbb{E}\left[\hat{\theta}_j^2\right] - \mathbb{E}\left[\hat{\theta}_j\right]^2 \\
&= \mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n}\left(f\left(X_i\right)+\epsilon_i\right)\varphi_j\left(X_i\right)\right)^2\right] - \left(\alpha_j+\theta_j\right)^2 \\
&= \mathbb{E}\left[\frac{1}{n^2}\sum_{i=1}^{n}\sum_{h=1}^{n}\left(f\left(X_i\right)+\epsilon_i\right)\varphi_j\left(X_i\right)\left(f\left(X_h\right)+\epsilon_h\right)\varphi_j\left(X_h\right)\right] - \left(\frac{1}{n}\sum_{i=1}^{n} f\left(i/n\right)\varphi_j\left(i/n\right)\right)^2 \\
&= \mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n} f\left(X_i\right)\varphi_j\left(X_i\right)\right)^2 + \frac{1}{n^2}\sum_{i=1}^{n}\epsilon_i^2\varphi_j\left(X_i\right)^2\right] - \left(\frac{1}{n}\sum_{i=1}^{n} f\left(i/n\right)\varphi_j\left(i/n\right)\right)^2 \quad \left(\mathbb{E}\left[\epsilon_i\right]=0\right) \\
&= \left(\frac{1}{n}\sum_{i=1}^{n} f\left(i/n\right)\varphi_j\left(i/n\right)\right)^2 + \frac{1}{n^2}\sum_{i=1}^{n}\mathbb{E}\left[\epsilon^2\varphi_j\left(i/n\right)^2\right] - \left(\frac{1}{n}\sum_{i=1}^{n} f\left(i/n\right)\varphi_j\left(i/n\right)\right)^2 \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2\varphi_j\left(i/n\right)^2 \\
&= \frac{\sigma^2}{n}\left[\frac{1}{n}\sum_{i=1}^{n}\varphi_j\left(i/n\right)\varphi_j\left(i/n\right)\right] \\
&= \frac{\sigma^2}{n}
\end{aligned}$$

(d)

$$\begin{aligned}
\mathbb{E}\left[\left(\hat{\theta}_j - \theta_j\right)^2\right] &= \mathrm{MSE}_{\theta_j}\left(\hat{\theta}_j\right) \\
&= \mathbb{E}\left[\hat{\theta}_j - \theta_j\right]^2 + \mathbb{E}\left[\left(\hat{\theta}_j - \mathbb{E}\left[\hat{\theta}_j\right]\right)^2\right] \\
&= \alpha_j^2 + \frac{\sigma^2}{n}
\end{aligned}$$

(e)

$$\begin{aligned}
\mathbb{E}\left[\int_0^1 \left(\hat{f}_{nN}(x) - f(x)\right)^2 dx\right] &= \mathbb{E}\left[\int_0^1 \left(\sum_{j=1}^{N} \hat{\theta}_j \varphi_j(x) - \sum_{j=1}^{\infty} \theta_j \varphi_j(x)\right)^2 dx\right] \\
&= \mathbb{E}\left[\int_0^1 \left(\sum_{j=1}^{N} \left(\hat{\theta}_j - \theta_j\right) \varphi_j(x) - \sum_{j=N+1}^{\infty} \theta_j \varphi_j(x)\right)^2 dx\right] \\
&= \int_0^1 \mathbb{E}\left[\left(\sum_{j=1}^{N} \left(\hat{\theta}_j - \theta_j\right) \varphi_j(x) - \sum_{j=N+1}^{\infty} \theta_j \varphi_j(x)\right)^2\right] dx \\
&= \int_0^1 \mathbb{E}\left[\left(\sum_{j=1}^{N} \left(\hat{\theta}_j - \theta_j\right) \varphi_j(x)\right)^2\right] \\
&\quad -2\left(\sum_{j=1}^{N} \mathbb{E}\left[\hat{\theta}_j - \theta_j\right] \varphi_j(x)\right)\left(\sum_{j=N+1}^{\infty} \theta_j \varphi_j(x)\right) + \left(\sum_{j=N+1}^{\infty} \theta_j \varphi_j(x)\right)^2 dx \\
&= \int_0^1 \mathbb{E}\left[\sum_{j=1}^{N}\sum_{k=1}^{N} \left(\hat{\theta}_j - \theta_j\right)\left(\hat{\theta}_k - \theta_k\right) \varphi_j(x)\varphi_k(x)\right] \\
&\quad -2\left(\sum_{j=1}^{N} \alpha_j \varphi_j(x)\right)\left(\sum_{j=N+1}^{\infty} \theta_j \varphi_j(x)\right) + \left(\sum_{j=N+1}^{\infty} \theta_j \varphi_j(x)\right)^2 dx \\
&= \mathbb{E}\left[\sum_{j=1}^{N}\sum_{k=1}^{N} \left(\hat{\theta}_j - \theta_j\right)\left(\hat{\theta}_k - \theta_k\right) \int_0^1 \varphi_j(x)\varphi_k(x)\,dx\right] \\
&\quad + \int_0^1 \left(\sum_{j=N+1}^{\infty} \theta_j \varphi_j(x)\right)^2 - 2\left(\sum_{j=1}^{N} \alpha_j \varphi_j(x)\right)\left(\sum_{j=N+1}^{\infty} \theta_j \varphi_j(x)\right) dx \\
&= \sum_{j=1}^{N} \mathbb{E}\left[\left(\hat{\theta}_j - \theta_j\right)^2\right] \\
&\quad + \sum_{j=N+1}^{\infty}\sum_{k=N+1}^{\infty} \theta_j\theta_k \int_0^1 \varphi_j(x)\varphi_k(x)\,dx - 2\sum_{j=1}^{N}\sum_{k=N+1}^{\infty} \alpha_j\theta_k \int_0^1 \varphi_j(x)\varphi_k(x)\,dx \\
&= \frac{\sigma^2 N}{n} + \sum_{j=1}^{N} \alpha_j^2 + \sum_{j=N+1}^{\infty} \theta_j^2
\end{aligned}$$

(f)

For all $n \geq 2$,

$$
\begin{aligned}
\max_{1\le j\le n-1}|\alpha_j| &= \max_{1\le j\le n-1}\left|\mathbb{E}\left[\hat{\theta}_j\right]-\theta_j\right| \\
&= \max_{1\le j\le n-1}\left|\frac{1}{n}\sum_{i=1}^{n} f\left(i/n\right)\varphi_j\left(i/n\right)-\theta_j\right| \\
&= \max_{1\le j\le n-1}\left|\frac{1}{n}\sum_{i=1}^{n}\left(\sum_{k=1}^{\infty}\theta_k\varphi_k\left(i/n\right)\right)\varphi_j\left(i/n\right)-\theta_j\right| \\
&= \max_{1\le j\le n-1}\left|\sum_{k=1}^{\infty}\theta_k\frac{1}{n}\sum_{i=1}^{n}\varphi_k\left(i/n\right)\varphi_j\left(i/n\right)-\theta_j\right| \\
&= \max_{1\le j\le n-1}\left|\sum_{k=1}^{n-1}\theta_k\frac{1}{n}\sum_{i=1}^{n}\varphi_k\left(i/n\right)\varphi_j\left(i/n\right)+\sum_{k=n}^{\infty}\theta_k\frac{1}{n}\sum_{i=1}^{n}\varphi_k\left(i/n\right)\varphi_j\left(i/n\right)-\theta_j\right| \\
&= \max_{1\le j\le n-1}\left|\sum_{k=1}^{n-1}\theta_k\delta_{kj}-\theta_j+\sum_{k=n}^{\infty}\theta_k\frac{1}{n}\sum_{i=1}^{n}\varphi_k\left(i/n\right)\varphi_j\left(i/n\right)\right| \\
&= \max_{1\le j\le n-1}\left|\sum_{k=n}^{\infty}\theta_k\frac{1}{n}\sum_{i=1}^{n}\varphi_k\left(i/n\right)\varphi_j\left(i/n\right)\right| \\
&\le \max_{1\le j\le n-1}\sum_{k=n}^{\infty}\left|\theta_k\frac{1}{n}\sum_{i=1}^{n}\varphi_k\left(i/n\right)\varphi_j\left(i/n\right)\right|
\end{aligned}
$$

Now observe that $|\varphi_k\left(i/n\right)\varphi_j\left(i/n\right)|\le 2$ for all k,j,i,n, which also implies $\left|\frac{1}{n}\sum_{i=1}^{n}\varphi_k\left(i/n\right)\varphi_j\left(i/n\right)\right|\le 2$. Hence

$$
\begin{aligned}
&\le \max_{1\le j\le n-1}\sum_{k=n}^{\infty}|\theta_k\left(2\right)| \\
&= 2\sum_{m=n}^{\infty}|\theta_m|
\end{aligned}
$$

(g)

We have that

$$\begin{aligned}
\sum_{m=n}^{\infty}|\theta_m| &\leq \sum_{m=n}^{\infty}\frac{a_m}{(m-1)^{\beta}}|\theta_m| \quad (\frac{a_m}{(m-1)^{\beta}}\geq 1)\\
&\leq \sqrt{\sum_{m=n}^{\infty}\frac{1}{(m-1)^{2\beta}}}\sqrt{\sum_{m=n}^{\infty}a_m^2\theta_m^2} \quad \text{(Cauchy-Schwarz inequality)}\\
&\leq \sqrt{\int_{m=n}^{\infty}\frac{1}{x^{2\beta}}dx}\sqrt{\sum_{m=n}^{\infty}a_m^2\theta_m^2} \quad \text{(integral is an upper bound to the sum)}\\
&= \sqrt{\left[\frac{x^{-2\beta+1}}{-2\beta+1}\right]_n^{\infty}}\sqrt{\sum_{m=n}^{\infty}a_m^2\theta_m^2}\\
&= \sqrt{\left[0-\frac{n^{-2\beta+1}}{-2\beta+1}\right]}\sqrt{\sum_{m=n}^{\infty}a_m^2\theta_m^2}\\
&= \frac{n^{-\beta+1/2}}{2\beta-1}\sqrt{\sum_{m=n}^{\infty}a_m^2\theta_m^2}\\
&\leq \frac{n^{-\beta+1/2}}{2\beta-1}\sqrt{\sum_{m=1}^{\infty}a_m^2\theta_m^2}\\
&\leq \frac{n^{-\beta+1/2}}{2\beta-1}\sqrt{Q}
\end{aligned}$$

and therefore

$$\begin{aligned}
\max_{1\leq j\leq n-1}|\alpha_j| &\leq 2\sum_{m=n}^{\infty}|\theta_m|\\
&\leq \frac{2\sqrt{Q}}{2\beta-1}n^{-\beta+1/2}\\
&= Cn^{-\beta+1/2}
\end{aligned}$$

where $C=\frac{2\sqrt{Q}}{2\beta-1}$.

(h)

We have that

$$\mathbb{E}\left[\int_0^1\left(\hat{f}_{nN}(x)-f(x)\right)^2dx\right] = \frac{\sigma^2N}{n}+\sum_{j=1}^{N}\alpha_j^2+\sum_{j=N+1}^{\infty}\theta_j^2$$

where

$$\begin{aligned}
\sum_{j=1}^{N} \alpha_j^2 &\leq N \max_{1 \leq j \leq N} \alpha_j^2 \\
&= N \left(\max_{1 \leq j \leq N} |\alpha_j| \right)^2 \\
&\leq N \left(C (N+1)^{-\beta+1/2} \right)^2 \quad \text{(from part (g))} \\
&= \mathrm{O}\left(\frac{C^2}{N^{2\beta-2}} \right)
\end{aligned}$$

and

$$\sum_{j=N+1}^{\infty} \theta_j^2 \leq \frac{Q}{N^{2\beta}}$$

because

$$\begin{aligned}
Q &\geq \sum_{m=1}^{\infty} a_m^2 \theta_m^2 \\
&= \sum_{m=1}^{\infty} (2m-2)^{2\beta} \theta_{2m-1}^2 + (2m)^{2\beta} \theta_{2m}^2 \\
&\geq \sum_{m=1}^{\infty} (2m-2)^{2\beta} \theta_{2m-1}^2 + (2m-1)^{2\beta} \theta_{2m}^2 \quad (\text{since } x^\beta \text{monotone, } \theta_{2m}^2 \geq 0) \\
&= \sum_{m=1}^{\infty} (m-1)^{2\beta} \theta_m^2 \\
&\geq \sum_{m=n}^{\infty} (m-1)^{2\beta} \theta_m^2 \\
&\geq (n-1)^{2\beta} \sum_{m=n}^{\infty} \theta_m^2 \\
\frac{Q}{(n-1)^{2\beta}} &\geq \sum_{m=n}^{\infty} \theta_m^2
\end{aligned}$$

Since the previous arguments apply to all $\theta \in \Theta(\beta, Q)$, they also apply to the supremum over θ. Therefore

$$\begin{aligned}
\sup_{\theta \in \Theta(\beta, Q)} \mathbb{E}\left[\int_0^1 \left(\hat{f}_{nN}(x) - f(x) \right)^2 dx \right] &\leq \frac{\sigma^2 N}{n} + \mathrm{O}\left(\frac{C^2}{N^{2\beta-2}} \right) + \frac{Q}{N^{2\beta}} \\
&= \frac{\sigma^2 \left\lfloor \alpha n^{\frac{1}{2\beta+1}} \right\rfloor}{n} + \mathrm{O}\left(\frac{C^2}{\left(\left\lfloor \alpha n^{\frac{1}{2\beta+1}} \right\rfloor \right)^{2\beta-2}} \right) + \frac{Q}{\left(\left\lfloor \alpha n^{\frac{1}{2\beta+1}} \right\rfloor \right)^{2\beta}} \\
&= \mathrm{O}\left(n^{-\frac{2\beta}{2\beta+1}} \right) + \mathrm{O}\left(n^{-\frac{2\beta-2}{2\beta+1}} \right) + \mathrm{O}\left(n^{-\frac{2\beta}{2\beta+1}} \right) \\
&= \mathrm{O}\left(n^{-\frac{2\beta}{2\beta+1}} \right) \\
&= C_0 n^{-\frac{2\beta}{2\beta+1}}
\end{aligned}$$

where C_0 depends on $\beta, Q, \alpha, \sigma^2$.

3

(a)

The mean of $\hat{\mu}_{bag}$ is

$$\begin{aligned}
\mathbb{E}\left[\hat{\mu}_{bag}\right] &= \mathbb{E}\left[\mathbb{E}\left[\hat{\mu}_{bag} | \{X_i\}_{i=1}^n\right]\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\frac{1}{2n}\sum_{i=1}^n Y_i^* + Z_i^* \mid \{X_i\}_{i=1}^n\right]\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[Y^* \mid \{X_i\}_{i=1}^n\right]\right] \quad (Y_i^*, Z_i^* \text{ are iid}) \\
&= \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n X_i\right] \\
&= \mathbb{E}\left[\bar{X}\right] \\
&= \mu
\end{aligned}$$

The variance of $\hat{\mu}_{bag}$ is

$$\begin{aligned}
\mathbb{V}\left[\hat{\mu}_{bag}\right] &= \mathbb{E}\left[\mathbb{V}\left[\hat{\mu}_{bag} \mid \{X_i\}_{i=1}^n\right]\right] + \mathbb{V}\left[\mathbb{E}\left[\hat{\mu}_{bag} \mid \{X_i\}_{i=1}^n\right]\right] \\
&= \mathbb{E}\left[\mathbb{V}\left[\frac{1}{2n}\sum_{i=1}^n Y_i^* + Z_i^* \mid \{X_i\}_{i=1}^n\right]\right] + \mathbb{V}\left[\bar{X}\right] \\
&= \mathbb{E}\left[\mathbb{V}\left[\frac{1}{2n}\sum_{i=1}^n Y_i^* + Z_i^* \mid \{X_i\}_{i=1}^n\right]\right] + \frac{\sigma^2}{n}
\end{aligned}$$

Since Y_i^*, Z_i^* are iid from $\{X_i\}_{i=1}^n$, let $U_i^* = Y_i^*$ for $i \in \{1, \ldots, n\}$ and $U_{n+i}^* = Z_i^*$ for $i \in \{1, \ldots, n\}$. Then

$$\begin{aligned}
&= \mathbb{E}\left[\mathbb{V}\left[\frac{1}{2n}\sum_{i=1}^{2n} U_i^* \mid \{X_i\}_{i=1}^n\right]\right] + \frac{\sigma^2}{n} \\
&= \mathbb{E}\left[\frac{1}{4n^2}\sum_{i=1}^{2n} \mathbb{V}\left[U^* \mid \{X_i\}_{i=1}^n\right]\right] + \frac{\sigma^2}{n} \\
&= \frac{1}{2n}\mathbb{E}\left[\mathbb{E}\left[(U^*)^2 \mid \{X_i\}_{i=1}^n\right] - \mathbb{E}\left[U^* \mid \{X_i\}_{i=1}^n\right]^2\right] + \frac{\sigma^2}{n} \\
&= \frac{1}{2n}\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^n X_i^2\right) - \bar{X}^2\right] + \frac{\sigma^2}{n} \\
&= \frac{1}{2n}\left(\mathbb{V}\left[X\right] + \mathbb{E}\left[X\right]^2 - \mathbb{V}\left[\bar{X}\right] - \mathbb{E}\left[\bar{X}\right]\right) + \frac{\sigma^2}{n} \\
&= \frac{\sigma^2 + \mu^2 - \frac{\sigma^2}{n} - \mu^2}{2n} + \frac{\sigma^2}{n} \\
&= \frac{(n-1)\,\sigma^2}{2n^2} + \frac{\sigma^2}{n}
\end{aligned}$$

Compare this to the variance of $\bar{X}$:

$$\begin{aligned}
\mathbb{V}\left[\bar{X}\right] &= \mathbb{V}\left[\frac{1}{n}\sum_{i=1}^n X_i\right] \\
&= \frac{1}{n^2}\sum_{i=1}^n \mathbb{V}\left[X\right] \\
&= \frac{\sigma^2}{n}
\end{aligned}$$

This is greater than $\mathbb{V}\left[\hat{\mu}_{bag}\right] = \frac{(n-1)\sigma^2}{2n^2} + \frac{\sigma^2}{n}$ by a positive term $\frac{(n-1)\sigma^2}{2n^2}$. Hence bagging does not improve the variance.

(b)

(1)

Noting that $\bar{X} \sim N\left(\mu, n^{-1}\right)$, the mean of $g(\mu)$ is

$$\begin{aligned}
\mathbb{E}\left[g(\mu)\right] &= \mathbb{E}\left[\mathbb{I}_{(\infty,\bar{X}]}(\mu)\right] \\
&= \int_{-\infty}^{\infty} \mathbb{I}_{(\infty,\bar{x}]}(\mu)\, f_{\bar{X}}(\bar{x})\, d\bar{x} \\
&= \int_{-\infty}^{\infty} \mathbb{I}_{(\infty,\bar{x}]}(\mu) \sqrt{\frac{n}{2\pi}} \exp\left\{-\frac{n(\bar{x}-\mu)^2}{2}\right\} d\bar{x} \\
&= \int_{\mu}^{\infty} \sqrt{\frac{n}{2\pi}} \exp\left\{-\frac{n(\bar{x}-\mu)^2}{2}\right\} d\bar{x} \\
&= \frac{1}{2}
\end{aligned}$$

The variance of $g(\mu)$ is

$$\begin{aligned}
\mathbb{E}\left[\left(g(\mu) - \frac{1}{2}\right)^2\right] &= \mathbb{E}\left[g(\mu)^2\right] - \mathbb{E}\left[g(\mu)\right]^2 \\
&= \mathbb{E}\left[\mathbb{I}_{(\infty,\bar{X}]}(\mu)^2\right] - \frac{1}{4} \\
&= \mathbb{E}\left[\mathbb{I}_{(\infty,\bar{X}]}(\mu)\right] - \frac{1}{4} \\
&= \frac{1}{2} - \frac{1}{4} \\
&= \frac{1}{4}
\end{aligned}$$

Both the mean and variance do not depend on n.

(2)

Noting that $\bar{X} \sim N\left(\mu, n^{-1}\right)$, the mean of $G(\mu)$ is

$$\begin{aligned}
\mathbb{E}\left[G(\mu)\right] &= \mathbb{E}\left[\Phi\left(\sqrt{n}\left(\mu - \bar{X}\right)\right) + o_P(1)\right] \\
&= \mathbb{E}\left[\Phi\left(\sqrt{n}\left(\mu - \bar{X}\right)\right)\right] + C_n \quad (\text{where } C_n \to 0 \text{ as } n \to \infty) \\
&= \int_{-\infty}^{\infty} \Phi\left(\sqrt{n}(\mu - \bar{x})\right) f_{\bar{X}}(\bar{x})\, d\bar{x} + C_n \\
&= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\sqrt{n}(\mu-\bar{x})} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{y^2}{2}\right\} dy\right] \sqrt{\frac{n}{2\pi}} \exp\left\{-\frac{n(\bar{x}-\mu)^2}{2}\right\} d\bar{x} + C_n
\end{aligned}$$

Observe that the integration is over an axis-aligned bivariate normal distribution with mean $(\mu, 0)$ in $(\bar{x}, y)$-space. The domain of integration is simply a half-plane passing through the mean $(\mu, 0)$, and which covers $\frac{1}{2}$ of $\mathbb{R}^2$. By the symmetry of the normal distribution about its mean,

$$\begin{aligned}
\mathbb{E}\left[G(\mu)\right] &= \frac{1}{2} + C_n \\
&\xrightarrow{n} \frac{1}{2}
\end{aligned}$$

4

(a)

Let h be the bayes classifier. The risk of Bayes rule is

$$R = \int P(Y \neq h(x)|X = x) f(x) dx$$

$$\begin{aligned}
P(Y \neq h(x)|X = x) &= P(Y = 0|X = x)P(m(x) > \frac{1}{2}) + P(Y = 1|X = x)P(m(x) < \frac{1}{2}) \\
&= \left[q + (1 - 2q) I\left(\sum_{j=1}^{J} x_j < \frac{J}{2} \right) \right] \cdot P\left(q + (1 - 2q) I\left(\sum_{j=1}^{J} x_j > \frac{J}{2} \right) > \frac{1}{2} \right) + \\
&\quad \left[q + (1 - 2q) I\left(\sum_{j=1}^{J} x_j > \frac{J}{2} \right) \right] \cdot P\left(q + (1 - 2q) I\left(\sum_{j=1}^{J} x_j > \frac{J}{2} \right) < \frac{1}{2} \right) \\
&= q \cdot P\left(I\left(\sum_{j=1}^{J} x_j > \frac{J}{2} \right) > \frac{1}{2} \right) + q \cdot P\left(I\left(\sum_{j=1}^{J} x_j > \frac{J}{2} \right) < \frac{1}{2} \right) \\
&= q
\end{aligned}$$

Thus, the risk of Bayes rule is q.

(b)

Prediction Performance:

	Density Tree	Naive Bayes	Kernel Regression	Additive Model	Random Forest
Error Rate	39.5%	46.5 %	38 %	35.5%	37 %