Supplementary Material for "Encompassing Tests for Nonparametric Regressions"

Elia Lapenta
CREST and ENSAE

Pascal Lavergne
Toulouse School of Economics

Proof of Lemma 8.3

In Part 1, we obtain the convergence rates of the derivatives of $\overline{Y}, \widehat{f}, \overline{\widehat{m}}, \overline{\phi_s(X)}, \overline{\xi\widehat{\varepsilon}}$, and $\overline{\widehat{m}^*}$. In Part 2, we use these rates to show the result of the Lemma.

Part 1. Fix a vector l of p positive integers with $|l| \leq \lceil (p+1)/2 \rceil$. For a differentiable function g, let $g^{(l)} := \partial^l g$. Since $\partial^l \overline{Y}(w) = h^{-(p+|l|)}\mathbb{P}_n Y K^{(l)}((w-W)/h)$, by Lemma 8.1 we have

$$\partial^l \overline{Y}(w) = \frac{1}{h^{p+|l|}} P Y K^{(l)}\left(\frac{w-W}{h}\right) + O_P\left(\sqrt{\frac{\log n}{n h^{p+2|l|}}}\right)$$

uniformly in $w \in \mathcal{W}$. For the leading term, we have (see the comments below)[1]

$$\begin{aligned}\frac{1}{h^{p+|l|}} P Y K^{(l)}\left(\frac{w-W}{h}\right) &= \frac{1}{h^p}\partial^l P\, Y\, K\left(\frac{w-W}{h}\right) = \frac{1}{h^p}\partial^l \int (mf)(\tilde{w}) K\left(\frac{w-\tilde{w}}{h}\right) d\tilde{w} \\ &= \partial^l \int (mf)(w+uh)\, K(u)\, du = \int (mf)^{(l)}(w+uh)\, K(u)\, du \\ &= \partial^l (mf)(w) + O(h^{r-|l|})\end{aligned}$$

uniformly in $w \in \mathcal{W}$. Indeed, by combining Assumption B(iii) and the Lebesgue Dominated Convergence Theorem, we can commute integration and differentiation to obtain the first equality. The second equality follows from the Law of Iterated Expectations, while the third one from a change of variable. The fourth equality commutes again integration and

[1] Notice that Assumption C(iii) implies that $r > \lceil (p+1)/2 \rceil$.

differentiation thanks to Assumption B, and finally the fifth one applies a Taylor expansion of order $r - |l|$. Putting together the previous displays gives

$$\|\partial^l \overline{Y} - \partial^l (mf)\|_\infty = O_P\left(d_n h^{-|l|}\right) \text{ for any } |l| \leq \lceil (p+1)/2 \rceil \,. \tag{S1}$$

An analogous reasoning gives

$$\|\partial^l \widehat{f} - \partial^l f\|_\infty = O_P(d_n h^{-|l|}) \text{ and } \sup_s \|\partial^l \overline{\phi_s(X)} - \partial^l (\iota_s f)\|_\infty = O_P\left(d_n h^{-|l|}\right) \tag{S2}$$

for any $|l| \leq \lceil (p+1)/2 \rceil$. We now show that

$$\|\partial^l \overline{\widehat{m}} - \partial^l (mf)\|_\infty = O_P\left(\frac{d_n}{h^{|l|}\tau_n} + \frac{p_n}{h^{p+|l|}}\right) \text{ for any } |l| \leq \lceil (p+1)/2 \rceil \,. \tag{S3}$$

To this end, notice first that

$$\begin{aligned}\partial^l \overline{\widehat{m}}(w) = & \frac{1}{h^{p+|l|}} \mathbb{P}_n m \, K^{(l)}\left(\frac{w-W}{h}\right) + \frac{1}{h^{p+|l|}} \mathbb{P}_n m(t-1) \, K^{(l)}\left(\frac{w-W}{h}\right) \\ & + \frac{1}{h^{p+|l|}} \mathbb{P}_n (\widehat{m} - m) t \, K^{(l)}\left(\frac{w-W}{h}\right) + \frac{1}{h^{p+|l|}} \mathbb{P}_n \widehat{m}(\widehat{t} - t) \, K^{(l)}\left(\frac{w-W}{h}\right) \,.\end{aligned} \tag{S4}$$

Let us handle each term separately. Since $\widehat{t} - t = (\widehat{t} + t)(\widehat{t} - t)$, $\|\widehat{m}(\widehat{t} + t)\|_\infty = O_P(1)$ from Lemma 8.1-(ii), $K^{(l)}$ is bounded, and $\mathbb{P}_n |\widehat{t} - t| = O_P(p_n)$, see the proof of Lemma 8.1-(i), the last term on the right-hand side is $O_P(p_n h^{-(p+|l|)})$ uniformly in $w \in \mathcal{W}$. Similarly, since $\mathbb{P}_n |t - 1| = O_P(p_n)$ from the proof of Lemma 8.1-(i), the second term on the right-hand side is $O_P(p_n h^{-(p+|l|)})$ uniformly in $w \in \mathcal{W}$. By reasoning as at the beginning of this proof, we get that the first term on the right-hand side is equal to $\partial^l (mf)(w) + O_P(d_n \, h^{-|l|})$ uniformly in $w \in \mathcal{W}$. Finally, the third term is bounded by

$$\frac{\|(\widehat{m} - m)t\|_\infty}{h^{|l|}} \frac{1}{h^p} \mathbb{P}_n \left| K^{(l)}\left(\frac{w-W}{h}\right)\right| = O_P\left(\frac{\|(\widehat{m} - m)t\|_\infty}{h^{|l|}}\right) = O_P\left(\frac{d_n}{h^{|l|}\tau_n}\right)$$

uniformly in $w \in \mathcal{W}$, where for the first equality we have used Lemma 8.1, while for the second equality we have used Equation (21) in Section 8. Gathering results gives (S3).

The convergence rate for $\partial^l \overline{\xi\widehat{\varepsilon}}$ is obtained by a similar reasoning, so

$$\|\partial^l \overline{\xi\widehat{\varepsilon}}\|_\infty = O_P\left(\frac{d_n}{h^{|l|}\tau_n} + \frac{p_n}{h^{p+|l|}}\right) \text{ for any } |l| \leq \lceil (p+1)/2 \rceil .$$

To get the convergence rate for $\partial^l \overline{\widehat{m}^*}$, we proceed as in (S4), where $\widehat{m}$ is replaced by $\widehat{m}^*$, and use the reasoning below it to obtain that

$$\|\partial^l \overline{\widehat{m}^*} - \partial^l (mf)\|_\infty = O_P\left(\frac{d_n}{h^{|l|}} + \frac{p_n}{h^{p+|l|}} + \frac{\|(\widehat{m}^* - m)t\|_\infty}{h^{|l|}}\right) .$$

By definition of $\widehat{m}^*$,

$$(\widehat{m}^* - m)t = \left(\frac{\overline{\widehat{m}}}{\widehat{f}} - m\right) t + \frac{\overline{\xi\widehat{\varepsilon}}}{\widehat{f}} t \,.$$

For the first term on the right-hand side we have that, uniformly over $\mathcal{W}$,

$$\left(\frac{\overline{\widehat{m}}}{\widehat{f}} - m\right) t = \frac{\overline{\widehat{m}} - m\widehat{f}}{f} t + \frac{\overline{\widehat{m}} - m\widehat{f}}{f} \frac{f - \widehat{f}}{\widehat{f}} t\, \mathbf{1}\left(\widehat{f}(\cdot) \geq \tau_n/2\right) = O_P\left(\frac{d_n}{\tau_n^2} + \frac{p_n}{h^p \tau_n}\right) .$$

Specifically, by Equation (19) evaluated at $\delta = 1$ we get that the first equality holds with probability approaching one. To obtain the second equality, we use $\|\widehat{f} - f\|_\infty = O_P(d_n)$ from the proof of Lemma 8.2-(ii) and $\|\overline{\widehat{m}} - mf\|_\infty = O_P\left(d_n \tau_n^{-1} + p_n h^{-p}\right)$ which follows from (S3) evaluated at $l = 0$. Using a similar reasoning, we can prove that uniformly over $\mathcal{W}$

$$\frac{\overline{\xi\widehat{\varepsilon}}}{\widehat{f}} t = O_P\left(\frac{d_n}{\tau_n^2} + \frac{p_n}{h^p \tau_n}\right) .$$

Gathering the previous four displays leads to

$$\|(\partial^l \overline{\widehat{m}^*} - \partial^l (mf)\|_\infty = O_P\left(\frac{d_n}{\tau_n^2 h^{|l|}} + \frac{p_n}{h^{p+|l|}\tau_n}\right) . \tag{S5}$$

Part 2. We now show that $\widehat{m} \in \mathcal{G}_l(\mathcal{W}_n)$ with probability approaching one (wpa1), for all l such that $|l| \leq \lceil (p+1)/2 \rceil$. By Assumption B there exists $\eta > 0$ such that $\|m^{(l)}\|_\infty + \eta < M$ for all $|l| \leq \lceil (p+1)/2 \rceil$. So, it suffices to show that wpa1 (a) $\widehat{m}$ is $\lceil (p+1)/2 \rceil$ times differentiable over $\mathcal{W}_n$, and (b) $\sup_w \left| \left[\widehat{m}^{(l)}(w) - m^{(l)}(w)\right] \mathbf{1}\left(f(w) \geq \tau/2\right) \right| \leq \eta$.

To show (a), we first use Equation (19) with $\delta = 1/2$ to get that

$$\text{wpa1 } \mathcal{W}_n \subset \left\{ w : \widehat{f}(w) \geq \tau_n/4 \right\} . \tag{S6}$$

Since by Assumption B(iii) $\widehat{m}$ is differentiable $\lceil (p+1)/2 \rceil$ times over $\{w : \widehat{f}(w) \geq \tau_n/4\}$, condition (a) follows.

To prove (b), define the event

$$\begin{aligned}
\mathcal{D}_n^C := & \left\{ \|(\widehat{m} - m)\mathbf{1}\left(f(\cdot) \geq \tau_n/2\right)\|_\infty \leq C d_n \tau_n^{-1} \right\} \\
& \cap \left\{ \|\overline{Y}^{(l)} - (mf)^{(l)}\|_\infty \leq C d_n h^{-|l|} \text{ and } \|\widehat{f}^{(l)} - f^{(l)}\|_\infty \leq C d_n h^{-|l|} \text{ for } |l| \leq \lceil (p+1)/2 \rceil \right\} \\
& \cap \left\{ \mathcal{W}_n \subset \{w : \widehat{f}(w) \geq \tau_n/4\} \right\} .
\end{aligned}$$

From Equation (21) evaluated at $\delta = 1/2$ and Equations (S1), (S2), and (S6), by choosing C large enough $P(\mathcal{D}_n^C)$ can be made arbitrarily close to 1 for each large n. Hence, to prove (b) it is sufficient to show that the event $\mathcal{D}_n^C$ implies that

$$\sup_{w \in \mathcal{W}_n} \left| \widehat{m}^{(l)}(w) - m^{(l)}(w) \right| \lesssim \frac{d_n}{\tau_n^{|l|} h^{|l|}} \text{ for } |l| \leq \lceil (p+1)/2 \rceil \tag{S7}$$

where the inequality holds up to a universal constant depending on C, M, and l.

Assume from now on that the event $\mathcal{D}_n^C$ holds. Then

$$\widehat{m}^{(l)} = \frac{\overline{Y}^{(l)}}{\widehat{f}} - \sum_{0 \leq |j| \leq |l|-1} \binom{l_1}{j_1} \cdots \binom{l_p}{j_p} \frac{\widehat{f}^{(l-j)}}{\widehat{f}} \widehat{m}^{(j)}$$

over $\mathcal{W}_n$. Similarly, over the set $\mathcal{W}_n$ we have

$$m^{(l)} = \frac{(mf)^{(l)}}{f} - \sum_{0 \leq |j| \leq |l|-1} \binom{l_1}{j_1} \cdots \binom{l_p}{j_p} \frac{f^{(l-j)}}{f} m^{(j)} .$$

Taking the difference between the previous two equations and recalling that the event $\mathcal{D}_n^C$

holds lead to

$$\sup_{w\in\mathcal{W}_n}\left|\widehat{m}^{(l)}(w)-m^{(l)}(w)\right| \le \frac{\|\overline{Y}^{(l)}-(mf)^{(l)}\|_\infty}{\tau_n} + \|(mf)^{(l)}\|_\infty \frac{\|\widehat{f}-f\|_\infty}{\tau_n^2}$$
$$+ \max_{0\le|j|\le|l|-1}\left\{\frac{||\widehat{f}^{(l-j)}-f^{(l-j)}||_\infty}{\tau_n}\left[\sup_{w\in\mathcal{W}_n}|\widehat{m}^{(j)}(w)-m^{(j)}(w)|+||m^{(j)}||_\infty\right]\right\}$$
$$+ \max_{0\le|j|\le|l|-1}\left\{\|f^{(l-j)}\|_\infty \frac{\sup_{w\in\mathcal{W}_n}|\widehat{m}^{(j)}(w)-m^{(j)}(w)|}{\tau_n}\right\}$$
$$+ \max_{0\le|j|\le|l|-1}\|f^{(l-j)}m^{(j)}\|_\infty \frac{\|\widehat{f}-f\|_\infty}{\tau_n^2}$$
$$\le \frac{d_n}{h^{|l|}\tau_n} + M^2\frac{d_n}{\tau_n^2} + \max_{0\le|j|\le|l|-1}\left\{\frac{d_n}{h^{|l|-|j|}\tau_n}\left[\sup_{w\in\mathcal{W}_n}|\widehat{m}^{(j)}(w)-m^{(j)}(w)|+M\right]\right\}$$
$$+ M\max_{0\le|j|\le|l|-1}\frac{\sup_{w\in\mathcal{W}_n}|\widehat{m}^{(j)}(w)-m^{(j)}(w)|}{\tau_n} + M^2\frac{d_n}{\tau_n^2} \tag{S8}$$

where the inequalities hold up to a universal constant that depends only on l and C.

Next, fix l such that $|l|=1$. By (S8)

$$\sup_{w\in\mathcal{W}_n}\left|\widehat{m}^{(l)}(w)-m^{(l)}(w)\right| \le \frac{d_n}{\tau_n h} + M^2\frac{d_n}{\tau_n^2} + \frac{d_n}{\tau_n h}\left[\frac{d_n}{\tau_n}+M\right] + M\frac{d_n}{\tau_n^2} + M^2\frac{d_n}{\tau_n^2} \le \frac{d_n}{\tau_n h}$$

up to a constant that depends only on M, C, and l, where in the last inequality we have used $h\tau_n^{-1}=o(1)$. Thus, (S7) holds for $|l|=1$.

We now proceed by induction. Fix l' such that $2\le|l'|\le\lceil(p+1)/2\rceil$ and assume that (S7) holds for any l such that $|l|\le|l'|-1$. Evaluating (S8) at l' leads to

$$\sup_{w\in\mathcal{W}_n}\left|\widehat{m}^{(l')}(w)-m^{(l')}(w)\right| \le \frac{d_n}{\tau_n h^{|l'|}} + M^2\frac{d_n}{\tau_n^2} + \max_{0\le|j|\le|l'|-1}\left\{\frac{d_n}{h^{|l'|-|j|}\tau_n}\left[\frac{d_n}{\tau_n^{|j|}h^{|j|}}+M\right]\right\}$$
$$+ M\max_{0\le|j|\le|l'|-1}\left\{\frac{d_n}{\tau_n^{|j|+1}h^{|j|}}\right\} + M^2\frac{d_n}{\tau_n^2} \le \frac{d_n}{h^{|l'|}\tau_n^{|l'|}}$$

where the inequalities hold up to universal constants that depend only on C, M, and l'. Since the induction assumption holds for $|l|=1$, we have proved that (S7) holds for all $|l|\le\lceil(p+1)/2\rceil$.

The proofs of (ii)-(v) follow from the same arguments and Equations (S2), (S3), and (S5).

This yields that for $|l| \leq \lceil (p+1)/2 \rceil$

$$\begin{aligned}
\sup_{w \in \mathcal{W}_n} \left| \partial^l \frac{\overline{\widehat{m}}}{\widehat{f}}(w) - \partial^l m(w) \right| &= O_P\left(\frac{d_n}{\tau_n^{1+|l|} h^{|l|}} + \frac{p_n}{h^{p+|l|} \tau_n^{|l|}} \right) \\
\sup_{s \in \mathcal{S}} \sup_{w \in \mathcal{W}_n} \left| \partial^l \widehat{\iota}_s(w) - \partial^l \iota_s(w) \right| &= O_P\left(\frac{d_n}{\tau_n^{|l|} h^{|l|}} \right) \\
\sup_{w \in \mathcal{W}_n} \left| \partial^l \frac{\overline{\xi \widehat{\varepsilon}}}{\widehat{f}}(w) \right| &= O_P\left(\frac{d_n}{\tau_n^{1+|l|} h^{|l|}} + \frac{p_n}{h^{p+|l|} \tau_n^{|l|}} \right) \\
\sup_{w \in \mathcal{W}_n} \left| \partial^l \frac{\overline{\widehat{m}^*}}{\widehat{f}}(w) - \partial^l m(w) \right| &= O_P\left(\frac{d_n}{\tau_n^{2+|l|} h^{|l|}} + \frac{p_n}{h^{p+|l|} \tau_n^{|l|+1}} \right) .
\end{aligned}$$