1 Nonparametric inference for the classical regression model

Consider the classical nonparametric regression model

$$Y_t = m(X_t) + e_t, \quad t = 1, \dots, T, \tag{1}$$

where Y_t, X_t and e_t are the responses, the predictors and the errors, respectively, and $m(\cdot)$ is an unknown smooth function. Suppose that X_t has compact support $\mathcal{X} \in \mathbb{R}$, then the trend function is defined as $m : \mathcal{X} \to \mathbb{R}$. For now, we consider scalar predictors X_t, however, for the future, the obvious generalisation would be to assume that $\mathcal{X} \in \mathbb{R}^d$ for some fixed d.

Let $K_X(\cdot)$ be a (potentially in the future d-dimensional) kernel function satisfying the following assumption:

(C1) The kernel K_X is non-negative, symmetric about zero and integrates to one. Moreover, it has compact support $[-1, 1]$ and is Lipschitz continuous, that is, $|K_X(v) - K_X(w)| \leq C|v - w|$ for any $v, w \in \mathbb{R}$ and some constant $C > 0$.

Consider a bandwidth h, a point $x \in \mathcal{X}$ and the corresponding kernel average

$$\widehat{\psi}_h(x) = \sum_{t=1}^{T} w_{t,h}(x) Y_t,$$

where $w_{t,h}(x)$ is a kernel weight defined at $x \in \mathcal{X}$. In order to avoid boundary issues, we work with a local linear weighting scheme. We in particular set

$$w_{t,h}(x) = \frac{\Lambda_{t,h}(x) K_X\left(\frac{X_t - x}{h}\right)}{\sum_{t=1}^{T} \Lambda_{t,h}(x) K_X\left(\frac{X_t - x}{h}\right)}, \tag{2}$$

where

$$\Lambda_{t,h}(x) = K_X\Big(\frac{X_t - x}{h}\Big)\Big[S_2(x) - \Big(\frac{X_t - x}{h}\Big) S_1(x)\Big],$$

and $S_\ell(x) = (Th)^{-1} \sum_{t=1}^{T} K_X(\frac{X_t - x}{h})(\frac{X_t - x}{h})^\ell$ for $\ell = 0, 1, 2$.

The kernel average $\widehat{\psi}_h(x)$ is nothing else than a rescaled local linear estimator of the function $m(\cdot)$ at a point x.

To allow nonstationary and dependent observations, we assume that the covariates X_t have the following properties (here t/T, $t = 1, \dots, T$, represents the time rescaled to the unit interval).

(C1) The variables X_t allow for the representation $X_t = H(t/T; \mathcal{G}_t)$, where $\mathcal{G}_t = (\dots, \xi_{t-1}, \xi_t)$, the random variables ξ_t are i.i.d. and $H : [0, 1] \times \mathbb{R}^\infty \to \mathcal{X}$ is a measurable function such that $H(t/T; \mathcal{G}_t)$ is well-defined for each t.

(C2) The value of $\mathbb{E}[H^2(t/T; \mathcal{G}_0)]$ is bounded away from zero and infinity on $[0, 1]$.

For the error process, we assume that

$$e_t = \sigma_t(X_t)\eta_t = \sigma(X_t, t/T)\eta_t,$$

where for now we consider i.i.d. η_t.
In order for the theory to work, we need the following assumptions: