Bosq,D., Blanke, D.*Inference and Prediction in Large Dimensions*. Wiley.

**Lemma 6.1** *(Bochner lemma)*
*If $H \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ is such that $\int H = 1$ and*

$$\| x \|^d \, |H(x)| \underset{\|x\| \to \infty}{\longrightarrow} 0,$$

*if $g \in L^1(\mathbb{R}^d)$, then for $H_\varepsilon(\cdot) := \varepsilon^{-d} H(\cdot/\varepsilon)$, one has $g * H_\varepsilon \underset{\varepsilon \to 0}{\longrightarrow} g$ at every continuity point of $g$.*

PROOF:
The condition $H \in L^\infty(\mathbb{R}^d)$ ensures that $g * H_\varepsilon$ is defined everywhere and we set $\Delta(x) = (g * H_\varepsilon)(x) - g(x)$. Since $\int H = \int H_\varepsilon = 1$, one may write $\Delta(x) = \int [g(x-y) - g(x)] H_\varepsilon(y) \mathrm{d}y$. First note that if $g$ is a bounded function (not necessarily in $L^1(\mathbb{R}^d)$!), the Bochner lemma easily follows from the dominated convergence theorem. In the general case, if $g$ is continuous at $x$, then given $\eta > 0$, we choose $\delta (= \delta_x) > 0$ such that for $\| y \| < \delta$, one has $|g(x-y) - g(x)| < \eta$. Next,

$$\begin{aligned} |\Delta(x)| &\le \eta \int_{\|y\| < \delta} |H_\varepsilon(y)| \mathrm{d}y + \int_{\|y\| \ge \delta} |g(x-y) - g(x)| |H_\varepsilon(y)| \mathrm{d}y \\ &\le \eta \, \| H \|_1 + \int_{\|y\| \ge \delta} |g(x-y)| |H_\varepsilon(y)| \mathrm{d}y + |g(x)| \int_{\|y\| \ge \delta} |H_\varepsilon(y)| \mathrm{d}y. \end{aligned}$$

Now since $H \in L^1$ and $\delta/\varepsilon \to \infty$ as $\varepsilon \to 0$, we have

$$\int_{\|y\| \ge \delta} |H_\varepsilon(y)| \mathrm{d}y = \int_{\|t\| > \delta/\varepsilon} |H(t)| \mathrm{d}t \underset{\varepsilon \to 0}{\longrightarrow} 0.$$

Moreover $g \in L^1(\mathbb{R}^d)$ and $\| t \|^d \, |H(t)| \to 0$ as $\| t \| \to \infty$ yield:

$$\int_{\|y\| \ge \delta} |g(x-y)| |H_\varepsilon(y)| \mathrm{d}y \le \delta^{-d} \sup_{\|t\| > \frac{\delta}{\varepsilon}} \| t \|^d \, |H(t)| \int_{\|y\| > \delta} |g(x-y)| \mathrm{d}y \to 0,$$

as $\varepsilon \to 0$ concluding the proof. ■

Coming back to density estimation, we begin with a necessary and sufficient condition for the $L^2$-consistency of the estimator (6.2).

**Theorem 6.1**
*The following assertions are equivalent:*

*(i)* $h_n \to 0$, $n h_n^d \to \infty$.
*(ii)* $\forall x \in \mathbb{R}^d$, *for every density $f$ continuous at $x$,* $\mathrm{E}(f_n(x) - f(x))^2 \underset{n \to \infty}{\longrightarrow} 0$.

PROOF: *(Outline)*

(1) (i) ⇒ (ii)

One may always write: $\mathrm{E}(f_n(x) - f(x))^2 = \mathrm{Var}\, f_n(x) + (\mathrm{E}\, f_n(x) - f(x))^2$. Now the Bochner lemma and $h_n \to 0$ imply that $\mathrm{E}\, f_n(x) \to f(x)$ and $n h_n^d \mathrm{Var}\, f_n(x) \to f(x) \, \| K \|_2^2$ as $n \to \infty$ at every continuity point of $f$. This relation, together with $n h_n^d \to \infty$, yields $\mathrm{Var}\, f_n(x) \to 0$.