K-NN Regression

General formula:

$$\sum_{j \in Q} \tilde{w}_j y(x_j)$$

$$Q \equiv \text{ the neighborhood of } x_0. \text{ (The } K \text{ nearest observations)}$$

$$\tilde{w}_i \equiv \frac{w_i}{\sum w}$$

Pre-Normalized Weight Function, w_{x_0,x_j}

→Must be a non-negative monotonically non-increasing function←

$w_{x_0,x_j} = \varphi(d) = \varphi(d_{x_0,x_j}) = \varphi(x_0, x_j)$, d is the distance between x_0 and x_j.

Shapes for $\varphi(d)$:

INFINITE WINDOW SIZE/infinite support

1. $\frac{1}{d+\epsilon}$, ϵ is something really small.
2. $\frac{1}{d^2+\epsilon}$
3. radial distance, $-d^2$ as exponent:
   - e^{-d^2}
   - $e^{-\frac{1}{2}d^2}$
   - $e^{-\frac{1}{2}\left(\frac{d}{\lambda}\right)^2}$ (Gaussian)

FINITE WINDOW SIZE 2λ/compact support (# of neighbors varies by observation)

4. $\varphi_\lambda(d) = \begin{cases} 1 - \left(\frac{d}{\lambda}\right)^2 \text{ if } 0 \le d \le \lambda \text{ and likewise } 0 \le \frac{d}{\lambda} \le 1 \\ 0, \text{ otherwise.} \end{cases}$
5. $\varphi_\lambda(d) = \begin{cases} \left(1 - \left|\frac{d}{\lambda}\right|^3\right)^3 \text{ if } 0 \le d \le \lambda \text{ and likewise } 0 \le \frac{d}{\lambda} \le 1 \\ 0, \text{ otherwise.} \end{cases}$
6. $\varphi_\lambda(d) = \begin{cases} 1 \text{ if } 0 \le d \le \lambda \\ 0, \text{ if } d > \lambda \end{cases}$

Trick: to keep the number of neighbors constant in the compact support cases allow λ to vary by observation (x_0):

$\lambda_{x_0,K}$ = distance to K^{th} nearest neighbor from x_0.

Kernels/Kernel Functions

1. Epanechnikov Kernel:

$$\varphi(x) = \begin{cases} 1 - x^2, \text{ if } -1 \le x \le 1 \text{ or } |x| \le 1 \text{ (an inverted parabola)} \\ 0, \text{ otherwise.} \end{cases}$$

   Note: For density estimation the area under a kernel must be 1, $\int\limits_{-\infty}^{\infty} \varphi(x)dx = 1$.

   To transform the Epanechnikov Kernel to a proper kernel density function note that:

$$\int\limits_{-\infty}^{\infty} \varphi(x)dx = \int\limits_{-1}^{1} (1 - x^2)dx = x - \frac{1}{3}x^3\Big|_{-1}^{1} = \frac{2}{3} - \left(-\frac{2}{3}\right) = \frac{4}{3}$$

   So, multiply $\varphi(x)$ by $\frac{3}{4}$ and you have a function that integrates to 1.

2. Tri-Cube Kernel:

$$\varphi(x) = \begin{cases} (1 - |x|^3)^3 \text{ if } -1 \le x \le 1 \\ 0, \text{ otherwise.} \end{cases}$$

   Likewise, one can confirm that $\int\limits_{-\infty}^{\infty} \varphi(x) = \frac{81}{70}$, so simply multiply the kernel by the reciprocal, $\frac{70}{81}$, and it is converted into a proper kernel density function.

3. Uniform ("Boxcar"/Rectangular) Kernel:

$$\varphi(x) = \begin{cases} 1 \text{ if } -1 \le x \le 1 \\ 0, \text{ otherwise.} \end{cases}$$

4. Gaussian Kernel: $e^{-\frac{1}{2}\left(\frac{x}{\lambda}\right)^2}$

   Using integration by parts, squaring, and converting to polar coordinates it can be shown that: $\int\limits_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{d}{\lambda}\right)^2} dx = \sqrt{2\pi} * \lambda$. So, multiply by $\frac{1}{\sqrt{2\pi} * \lambda}$ to get a proper density function.