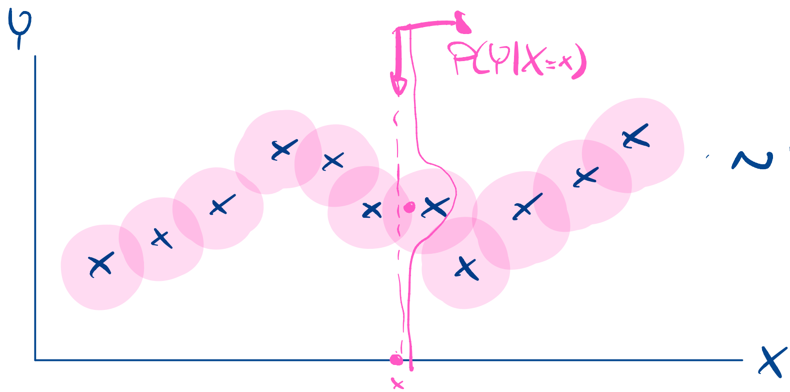
NADARAYA - WATSON

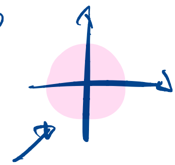
Dataset $\{(x_i, y_i)\}_{i=1}^{N}$



$\sim P(X,Y)$

base densities

$f(x)\, g(y)$ where f, g are unimodal, localized with mean 0



(think of $\dfrac{e^{-(x^2+y^2)/(2\sigma^2)}}{2\pi\sigma^2}$)

model for the joint density

$$p(x,y) := \frac{1}{N}\sum_{i=1}^{N} f(x-x_i)\, g(y-y_i)$$

$$\int p(x,y)\, dx\, dy = \frac{1}{N}\sum_{i=1}^{N} \underbrace{\int f(x-x_i)\, dx}_{\overset{\tilde{x} := x - x_i}{=} \int f(\tilde{x})\, d\tilde{x} = 1} \; \underbrace{\int g(y-y_i)\, dy}_{=1}$$

$$= 1$$

prediction $\quad y_{model}(x) := \mathbb{E}_{\mathbb{P}(Y|X=x)}[Y] = \int \frac{p(x,y)\, y}{\int p(x,\tilde{y})\, d\tilde{y}}\, dy$

$$= \sum_{i=1}^{N} f(x-x_i) \int g(y-y_i)\, y\, dy \; \frac{1}{\sum_{j=1}^{N} f(x-x_j) \underbrace{\int g(\underbrace{\tilde{y}-y_j}_{=y})\, d\tilde{y}}_{=1}}$$

$y' := y - y_i$

$$= \sum_{i=1}^{N} \underbrace{\frac{f(x-x_i)}{\sum_{j=1}^{N} f(x-x_j)}}_{=: k(x,x_i)} \underbrace{\int g(y')\,(y'+y_i)\, dy'}_{\underbrace{\int g(y')\, y'\, dy'}_{=0} + y_i \underbrace{\int g(y')\, dy'}_{=1}}$$

$$= \underline{\sum_{i=1}^{N} y_i\, k(x,x_i)}$$

- $\sum_{i=1}^{N} k(x,x_i) = \frac{\cancel{\sum_{i=1}^{N} f(x-x_i)}}{\cancel{\sum_{j=1}^{N} f(x-x_j)}} = 1$
- $k(x,x_i) = \frac{f(x-x_i)}{f(x-x_i) + \sum_{j\neq i} f(x-x_j)} = \frac{1}{1+\ldots} \leq 1$

$\Big\}$ $k(x,x_i)$ is a prob. dist. over i

$\Rightarrow y_{model}(x) =$ expectation of $\underline{y_i}$ wrt $i \mapsto \underline{k(x,x_i)}$

$$k(x, x_i) = \frac{\overbrace{f(x - x_i)}^{\text{large at } x \approx x_i}}{\underbrace{\sum_{j=1}^{N} f(x - x_j)}_{\text{more or less const wrt } x}}$$