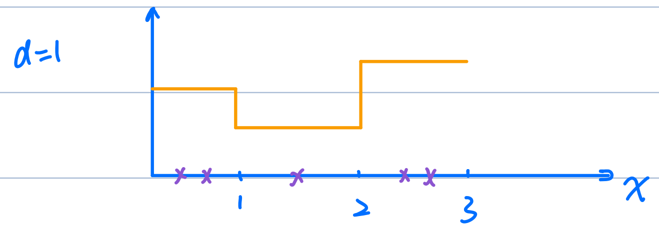
1. Recall: density estimation:

↳ Given $\mathcal{D} = \{ \underline{x}_1 \dots \underline{x}_N \}$ $\quad x_i \overset{iid}{\sim} p(\underline{x})$ unknown

Find $\hat{p}(\underline{x}) = p(\underline{x})$

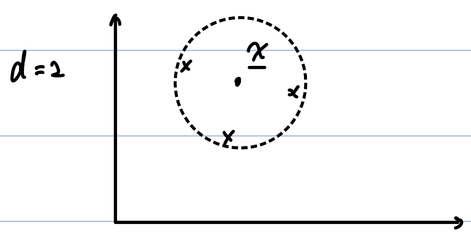
↳ Histogram Method:



Problem: Not continuous. ∴ GD may have problem

Sol: Nearest Neighbor estimation

2. Nearest Neighbor estimation



↳ For any $\underline{x} \in \mathbb{R}^d$ (Not near a sample point)

Let $\underline{x}_{[1]}, \dots \underline{x}_{[k]}$ be the nearest k neighbors of $\underline{x}$, sorted in ↑ distance

Let $r_k(\underline{x}) = \| \underline{x} - \underline{x}_{[k]} \|$

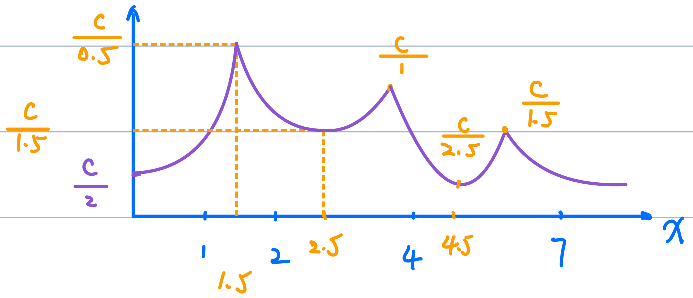
$V_k(\underline{x})$ = volume of sphere in $\mathbb{R}^d$ with radius $r_k(\underline{x})$

Then $\hat{p}(\underline{x}) \propto \frac{1}{V_k(\underline{x})}$

i.e. $\hat{p}(\underline{x}) = \frac{c}{V_k(\underline{x})}$, where c to be determined

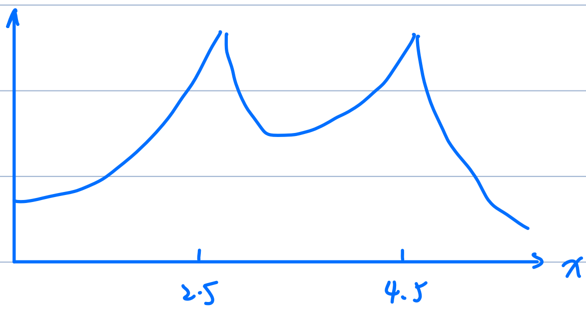
↳ To find c, set $\int \hat{p}(\underline{x}) = 1$

Example:



$\mathcal{D} = \{1, 2, 4, 7\}$, $k = 2$

d=3



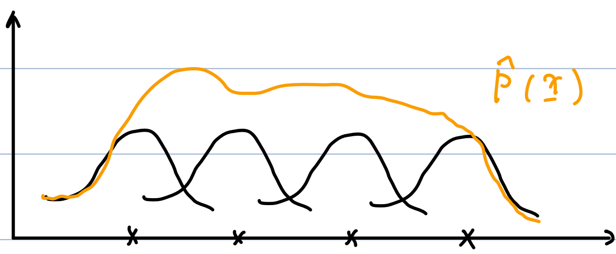
Larger k ⇒ smoother output curve

if k=1 ⇒ delta function

3. Parzen Window Estimation with Guassian kernels:

(1962) (Rosenblaff 1956)

↳ Idea:



↳ Details:

Standard normal distribution (pdf)

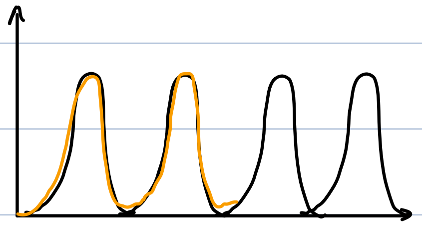
$$\phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)$$

$$\hat{p}(\underline{x}) = \frac{1}{K} \sum_{i=1}^{N} \phi\left(\frac{\|\underline{x} - \underline{x}_i\|}{\gamma}\right) ,$$

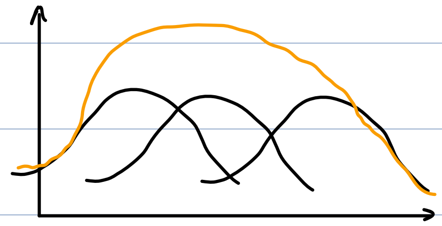
// $\underline{x} \in \mathbb{R}^d$, γ: kernel width, K: Normalizing constant

↳ Effect of Different γ :

small γ — large γ



over fitting



under fitting

↳ Problem:

Many sample points ⇒ hard to plot