Probabilistic Prediction

- To handle ambiguous cases

Probability Distribution : $p(y \mid x, D)$

= a Vector of length C ;

- If $C = 2$; $p(y=1 \mid \underline{x}, D) + p(y=0 \mid \underline{x}, D) = 1$

- $\hat{y} = \hat{f}(\underline{x}) = \underset{c=1}{\overset{C}{\operatorname{argmax}}}\ P(y = c \mid \underline{x}, D)$

~ most probable class label

Regression

- Response variable is continuous

$x_i \in \mathbb{R} \rightarrow y_i \in \mathbb{R}$

Unsupervised learning

~ density estimation i.e. Build a model

; unconditional density estimation.

a model of form $P(\underset{\downarrow}{x_i} \mid \theta)$;

vector of features

So we need to create multivariate prob. models.

- First task is to estimate the distributions over the number of clusters $p(K \mid D) \Rightarrow$ any subpopulation there are not

- for dimensionality reduction we use principal components analysis or PCA :

① Parametric vs non-parametric model

probabilistic model are of form $p(y|\underline{x})$ or $p(\underline{x})$

$p(y|\underline{x}) \rightarrow$ supervised, $p(\underline{x}) \rightarrow$ U.S

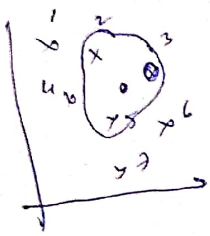
faster $\leftarrow$ (Parametric) parameters of the model is fixed ?

$\therefore$ flexible $\leftarrow$ (Non parametric) grows or with the amount of traing data ?

Ex (Non parametric): KNN = $p(y=c|\underline{x}, D, K) = \frac{1}{K} \sum_{i \in N_K(\underline{x}, D)} \mathbb{I}(y_i = c)$

$N_K(\underline{x}, D)$ = indices of K nearest points to $\underline{x}$ in D

$\mathbb{I}(e)$ = indicator funct^n = $\begin{cases} 1 & \text{if } e = \text{true} \\ 0 & \text{if } e \text{ is false} \end{cases}$



$p(y=0|\underline{x}, D, K) = \frac{1}{K} \sum_{i \in N_K} \mathbb{I}(y_i = 0)$.

$N_K(\underline{x}, D) = \{2, 3, 5\}$

$= \frac{1}{3}\{0, 1, 0\} = \frac{1}{3}$

$p(y=1|\underline{x}, D, K) = \frac{1}{K} \sum_i \quad \frac{1}{3}\{1, 0, 1\} = 2/3.$

② Linear Regression :

response is linear funct^n of input

$y(\underline{x}) = \underline{w}^T \underline{x} + \varepsilon = \sum_{j=1}^{D} w_j x_j + \varepsilon$

$\varepsilon \rightarrow$ residual Error

Gaussian or Normal distribution $N(\mu, \sigma^2)$