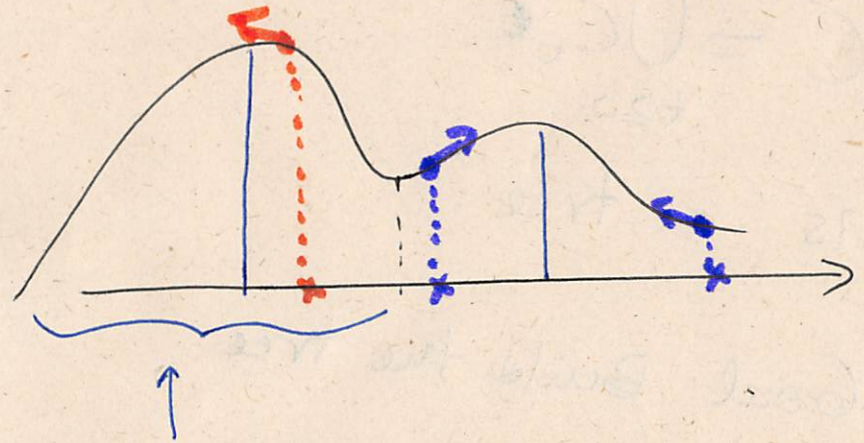
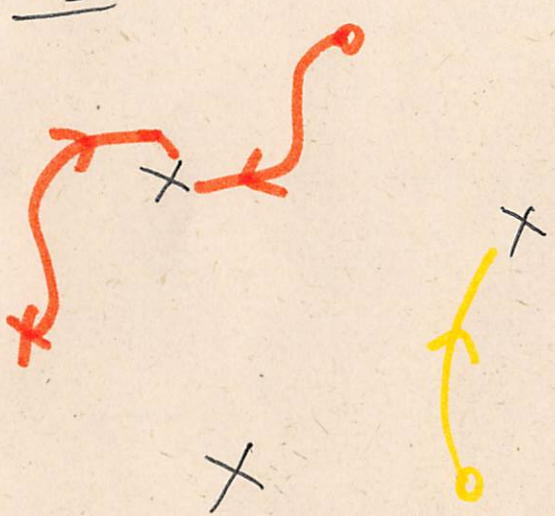
Mode = local max.

1D:



basin of attraction

2D:



## Define Paths:

$$\Pi_x(t)$$

$$\begin{cases} \Pi_x(0) = 0 \\ \Pi_x'(t) = \nabla_p(\Pi_x(t)) \end{cases}$$

$$\text{dest}(x) = \lim_{t \to \infty} \Pi_x(t) \in \text{set of modes}$$

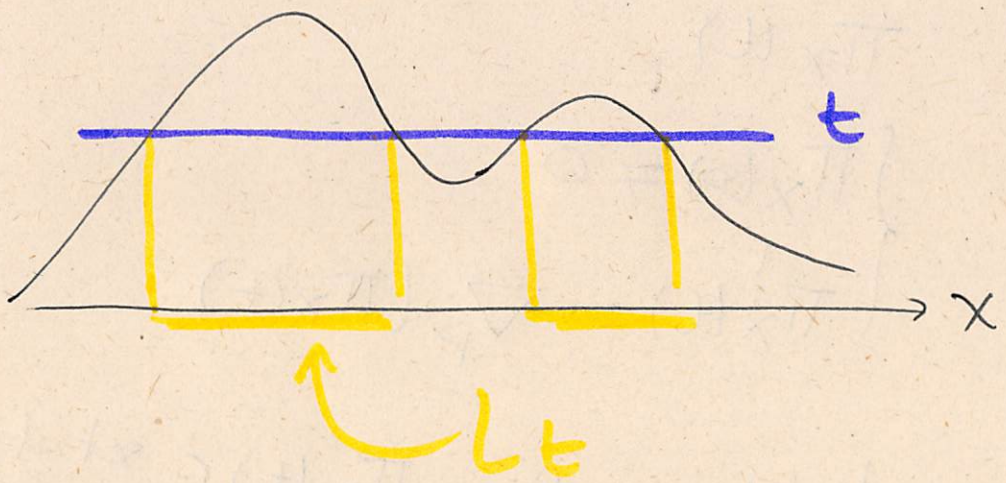
## MS-Algo

⓪ Pick any point $a = a^{(0)}$.

①

$$a^{(1)} \leftarrow \frac{\sum_{i=1}^{n} x_i K\left(\frac{x_i - a_k^{(0)}}{h}\right)}{\sum_{i=1}^{n} K\left(\frac{x_i - a_k^{(0)}}{h}\right)}$$

# Density tree




$$L_t := \{x : p(x) > t\} = \bigcup_j C_j$$

- High density region is an intuitive notion of cluster

$$C = \bigcup_{t \geq 0} C_t$$

is a tree.

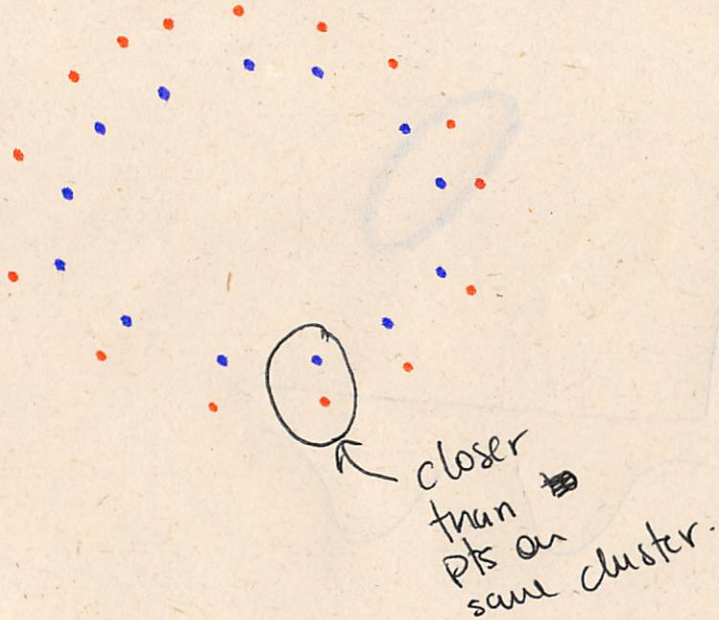
Goal: Build the tree.

$X \sim N(0, I)$ ; $X \in \mathbb{R}^d$

Not hard to show that:

$$\|x\|^2 = d$$

i.e. points lie on rings.




One fix:

Modify the distance

$$\rho(x, y) := \frac{1}{n-2} \sum_{z \neq x, y} \left| \|x - z\| - \|y - z\| \right|$$

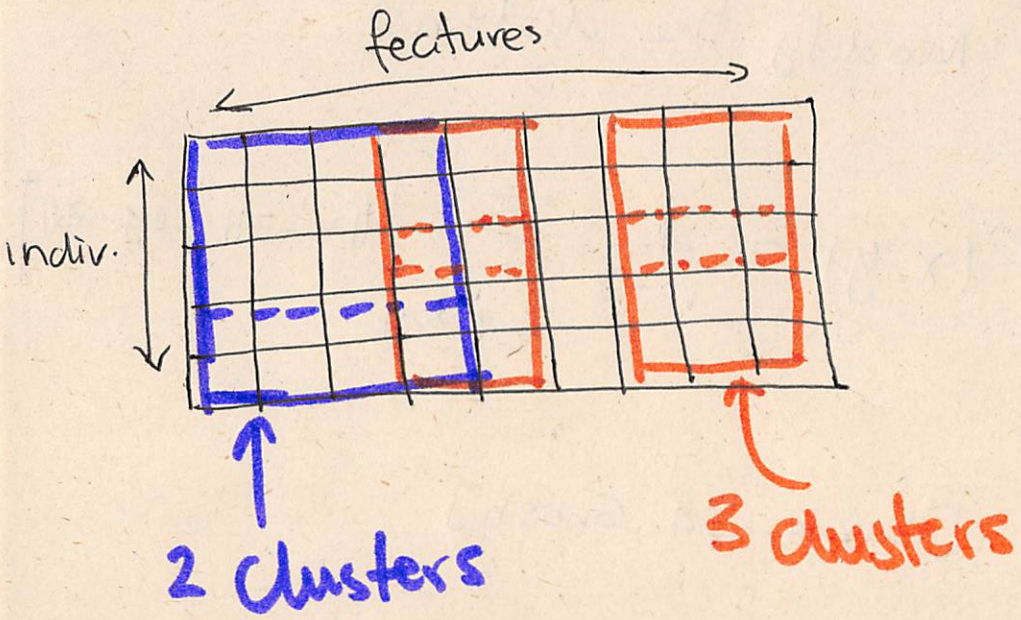
[Sarkar & Ghosh]

⚠ Proofs assume independence between coordinates...

Open problem:

Adjust k-means to deal with high dim. effects

Find Features that define clusters.




(interesting cluster)

## A First Idea:

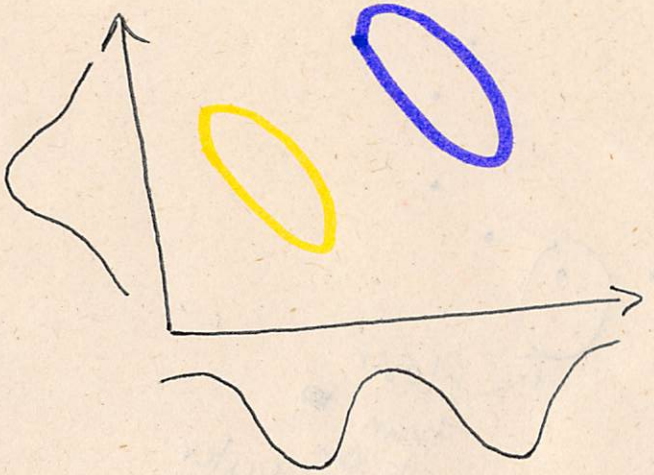
Screening: Are there marginal clusters?

$X = (X[1], \ldots, X[n])$.

$X_1[1], \ldots, X_m[n] \sim F$

$H_0$: $F$ is unimodal.

$H_1$: $F$ is <u>not</u>.

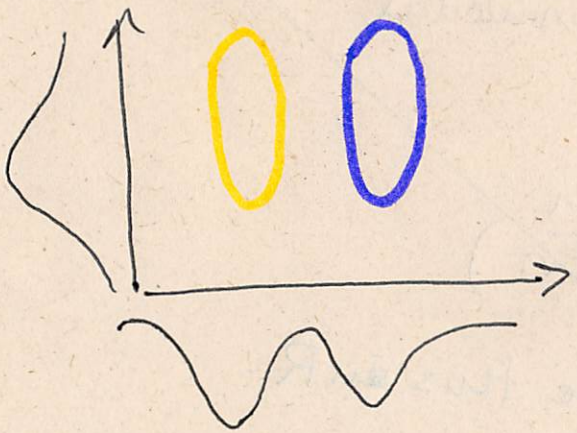


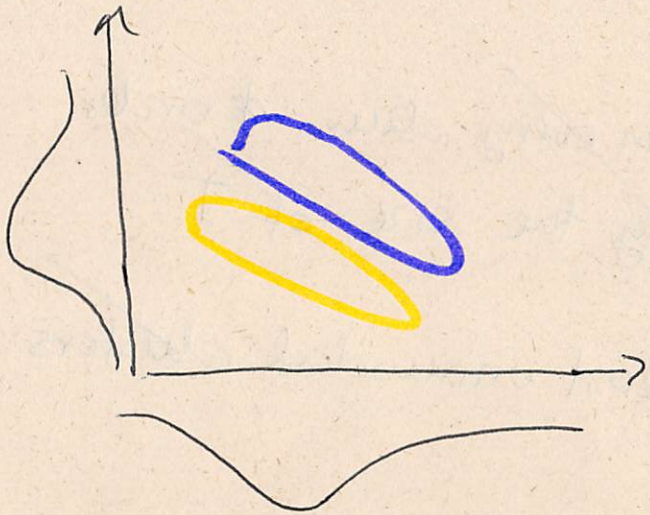
$$\text{Dip}(F) = \inf_{G \text{ unimodal}} \sup_x |F(x) - G(x)|$$

$$T := \text{Dip}(\hat{F}_n)$$

↑ can compute this in R

- Do this in every dim. & order the T's.
- Select least unimodal?





$DIP(F) := \inf_{G \text{ unimodal}} \sup_x |F(x) - G(x)|$

Sparse Alternating Clustering

$T = DIP(\hat{F}_n)$

↑ can compute this in R.

- Do this in every dim. & order them by the size of T.
- select least unimodal clusters.

$$R_n = \frac{1}{n} \sum_{j=1}^{n} \min_{c} \|x_j - c\|^2$$

$$= \dots =$$

$$= \sum_{j=1}^{k} \frac{1}{|C_j|} \sum_{i,j} \|x_i - x_j\|^2$$

Choose $|S| = L$.

$a = 1, \dots, d$ index of features

$$\delta_a(i,j) := (x_i[a] - x_j[a])^2$$

Rescale so that $\sum_{i,j} \delta(i,j) = 1$

Do k-means with an alternative step:

- Find the centers.
- Find the features

↑

- minimize $R_n$ to find the features.