

MPE

ANOVA

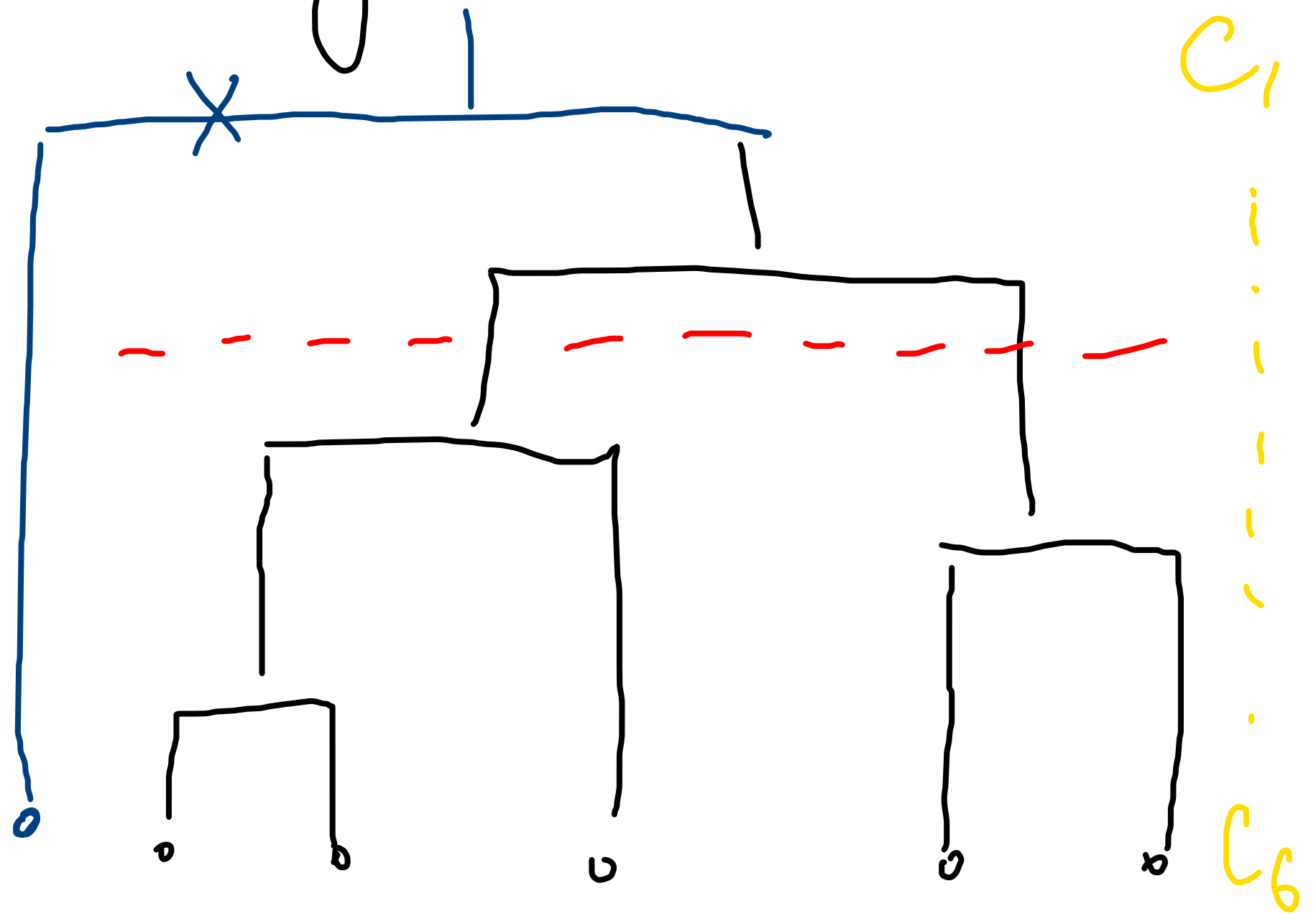
F

$=$

$$\frac{\text{Var Between}}{\text{Var Within}}$$

Partition clustering vs Hierarchical

- k-means



$K = 2$; Complete linkage

$\delta(x_i, x_j)$	1	2	3	4	5
1	0	8	8	7	6
2		0	2	4	5
3			0	3	2
4				0	1
5					0

$K = 5$

C.L.

$\Rightarrow 7$

S.L.

$\Rightarrow 6$

A.L.

$\Rightarrow 6, 5$

a_j

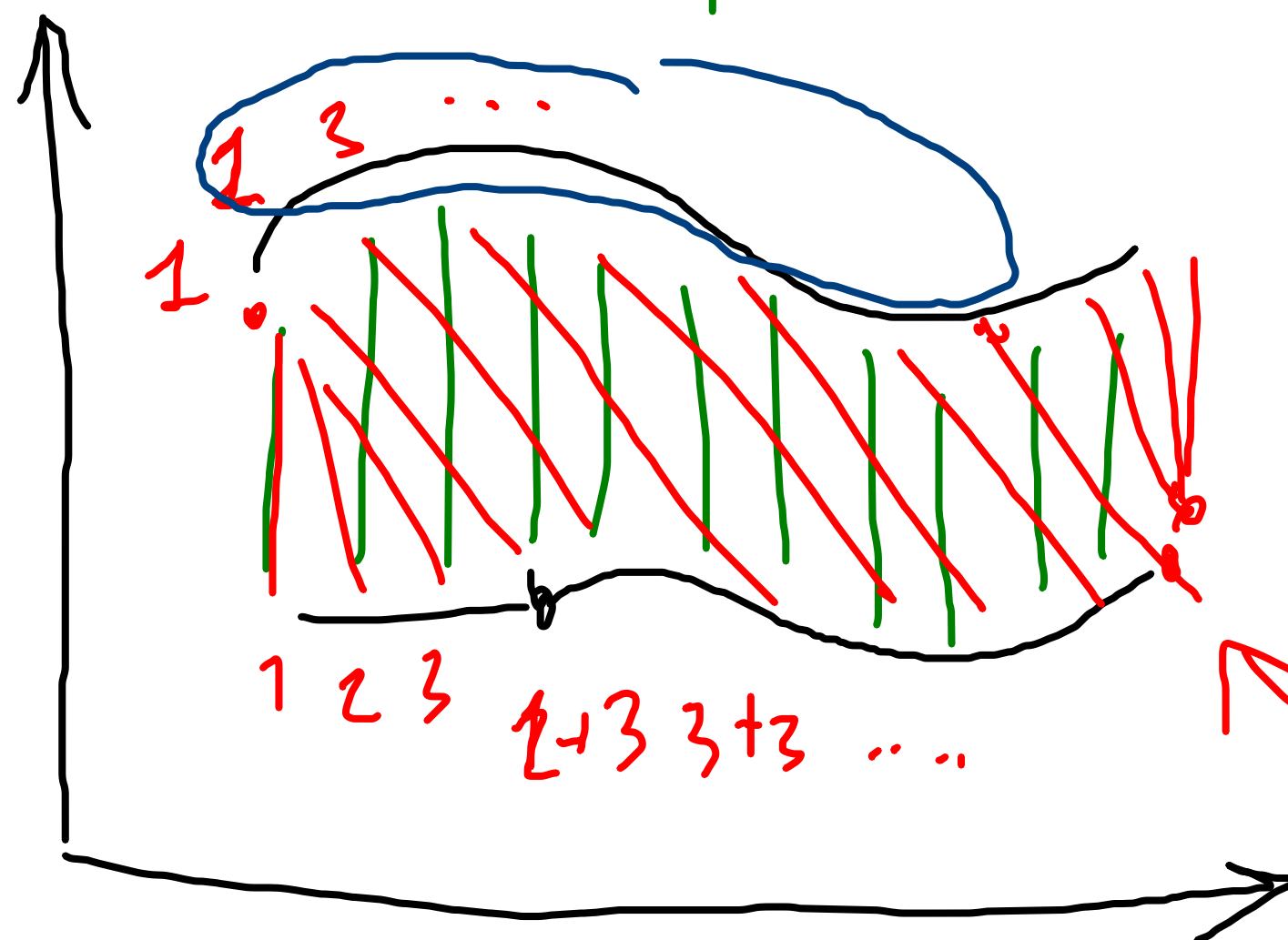
$$\delta(x_1, x_2) = \sqrt{\sum (x_{1i} - x_{2i})^2}$$

TS:

$$\delta(x_1, x_2) = 1 - \hat{p}_{x_1, x_2}$$

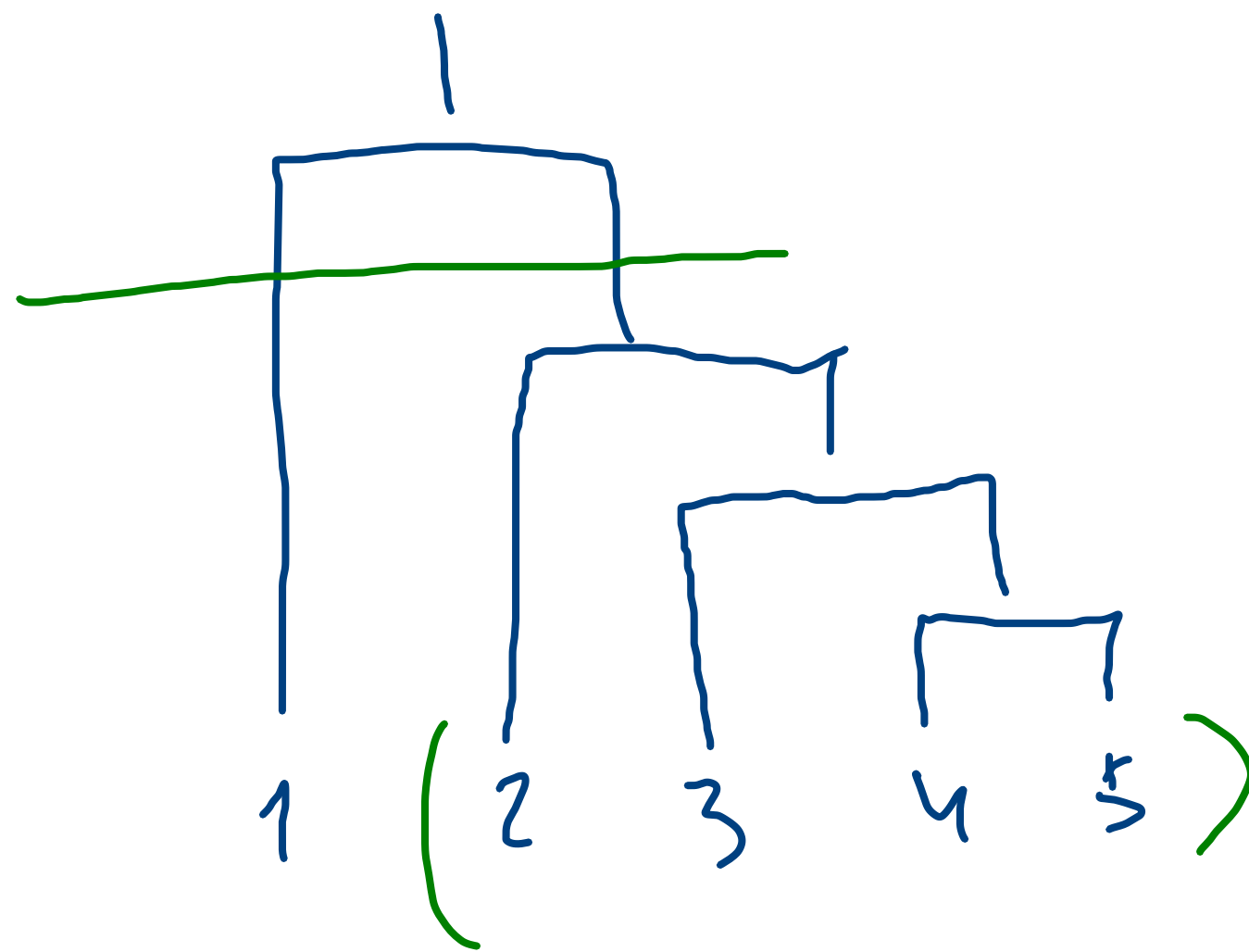
DTW

eqg.



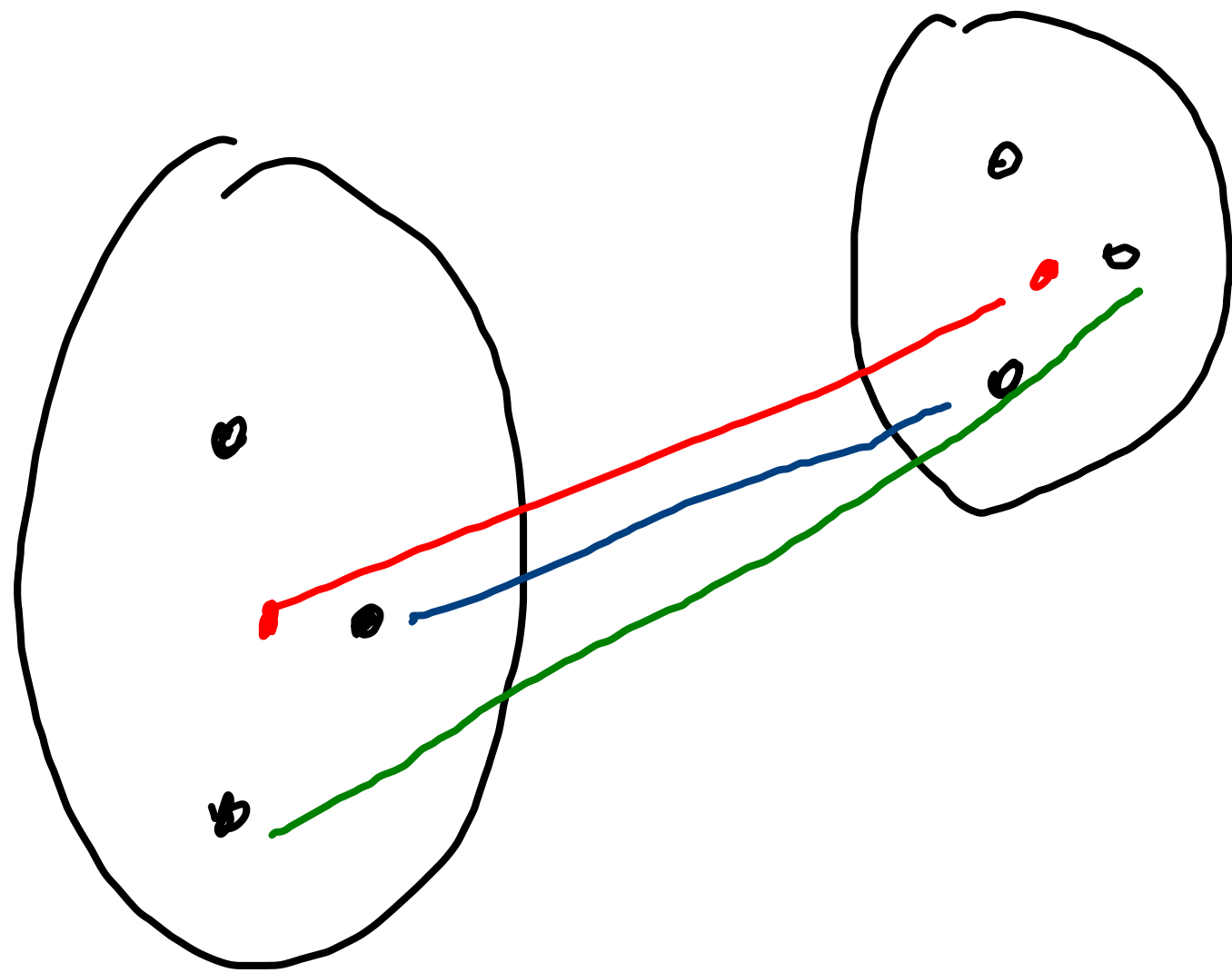
$$i_{t-1} \leq i_t \leq i_{t+1}$$

$$j_{t-1} \leq j_t \leq j_{t+1}$$



	1	2	13, 4, 5)
1	0	8	8
2		0	7
13, 4, 5)			0

	1	2	3	(4, 5)
1	0	8	8	7
2		0	7	5
3			6	3
(4, 5)				0



$$\sum (x_i - \bar{x})^2$$

$$5) \delta SSE_{ij} =$$

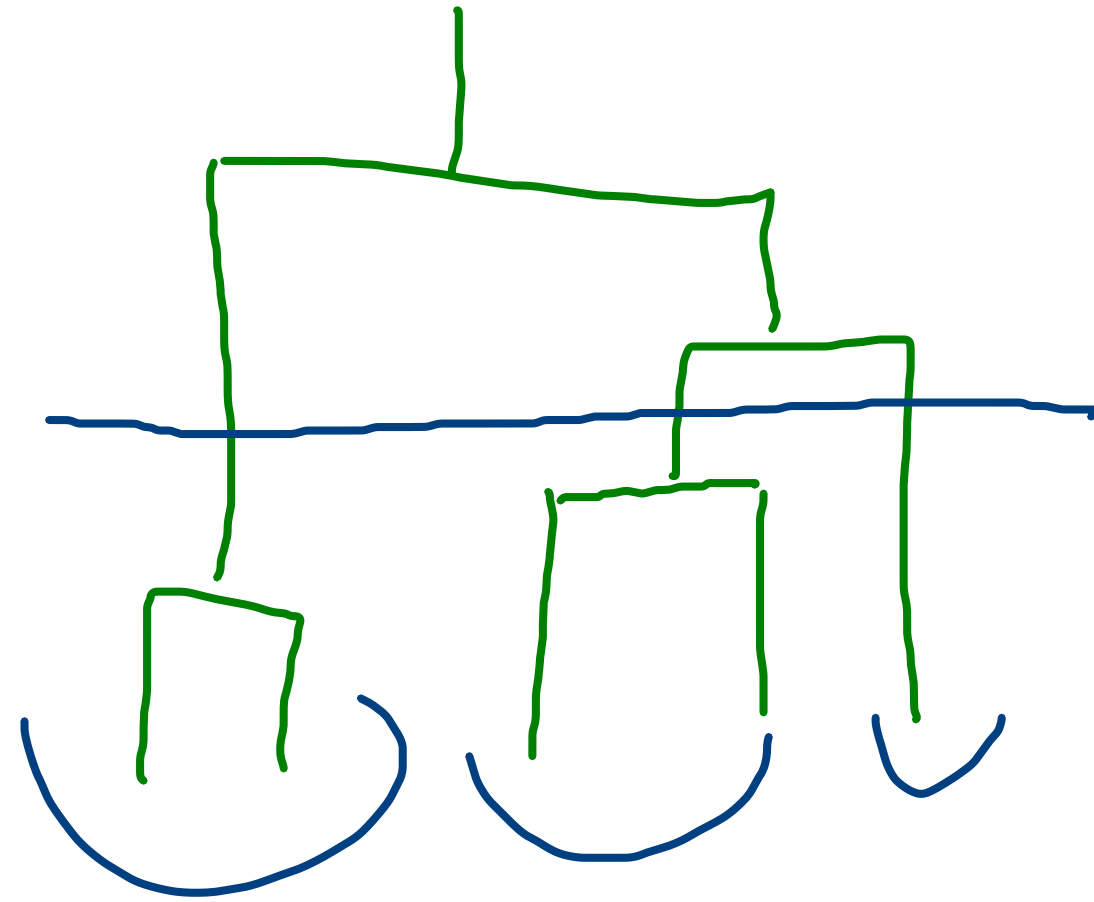
$$= SSE_{ij} - SSE_i - SSE_j$$

1) Compare centroids : mean / median

2) Single linkage $\min \delta_{ij}$

3) Complete linkage $\max \delta_{ij}$

4) Average linkage $\text{mean } \delta_{ij}$



$$y_i = x_i^T \beta + \varepsilon_i$$

1) linearity + correct. specification

2) $E(\varepsilon_i | X) = 0$
 $\hookrightarrow \text{cov}(\varepsilon_i, x_i) = 0$ No {
 - omitted var.
 - measurement er.
 - simultaneity

3) Homogeneity of obj: $E(\varepsilon_i^2) = \text{Var}(\varepsilon_i^2) = \sigma_\varepsilon^2$

4) $E(\varepsilon_i, \varepsilon_j) = 0$

5) $\text{rank}(X) = k+1$