

# Lecture 28

## Kernel Clustering with Quantum Circuits

*of the course “Fundamentals of Quantum Computing”*

(by  and **QUANTERALL**)

Stoyan Mishev, Vesselin Gueorguiev and Vladimir Gerdjikov



INSTITUTE *for* ADVANCED  
PHYSICAL STUDIES



February 24, 2023



$$i=5 \rightarrow z_i=2$$

$$i=105 \rightarrow z_i=1$$

Cost function

K-means

data points

cluster centroids

$$C(z, \mu) = \sum_i \|x_i - \mu_{z_i}\|^2$$

index of the cluster  $x_i$  belongs to

The task is to find  $z_i$  and  $\mu_{z_i}$ :  $C(z, \mu)$  is minimal

Iterative Algorithm:

$z_i = \arg \min_k \|x_i - \mu_k\|^2$  (for  $x_i \rightarrow z_i$ )

$$\mu_k = \frac{1}{N_k} \sum_{i \in C_k} x_i$$

# of points belonging to the cluster  $k$

$\langle \varphi(x_i), \varphi(x_i) \rangle$  Feature map  $\varphi(x_i)$

$K$  clusters

Cost:  $C(\mu) = \sum \|\varphi(x_i) - \mu_k\|^2$

if  $x_i \rightarrow X_N$  then  $\varphi(x_i) \rightarrow \varphi(X_N)$  def  $K \times N$

Define  $Z$

|   |   |   |   |   |
|---|---|---|---|---|
| 0 | 0 | 1 | 0 | 0 |
| 0 | 1 | 0 | 0 | 0 |
| 0 | 0 | 0 | 1 | 0 |
| 1 | 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 | 0 |

$N \times K$  - # of cluster

- assignment matrix  
- \* row shows which point belongs to which cluster

$\sum_k z_{nk} = 1$

$\sum_n z_{nk} = N_k$

$\rightarrow L = \text{diag}(1/N_k) = \text{diag}(1/\sum_n z_{nk}) \rightarrow K \times K$  matrix

Construct

$\varphi(X_N) Z L Z^T = M$

$K \times N$   $N \times K$   $K \times K$   $K \times N$   $K \times N$

$K$  { 5 5 2  
5 5 3  
1 1 1  
1 1 1  
2 2 2 }

\* column contains a copy of the cluster mean assigned to corresp. data point

The objective can be rewritten as:  
(cost function)

$$J = \Phi Z L Z^T$$

$$\rightarrow C = \text{tr} \left[ \left( \Phi_{K \times N} - \tilde{M}_{K \times N} \right) \left( \Phi_{K \times N} - \tilde{M}_{K \times N} \right)^T \right]$$

Proof that  $(Z L Z^T)^2 = Z L Z^T$  (projector)  $\Rightarrow (I - Z L Z^T)^2 = I - Z L Z^T$

$$\Rightarrow C = \text{tr}(\Phi \Phi^T) - \text{tr}(\Phi Z L Z^T \Phi^T) =$$

K - kernel matrix

$$= \text{tr}(K) - \text{tr}(L^{1/2} Z^T K Z L^{1/2})$$

The task is to binary matrix  $Z$ :  $\max_{H^T} \text{tr}(L^{1/2} Z^T K Z L^{1/2})$

✓ (this is hard)

$\hookrightarrow H = Z L^{1/2}$  is not binary anymore and

the task now is to find  $\max_H \text{tr}[H^T K H]$ :  $H^T \cdot H = I$

If we interpret the columns of  $H$  as a collection of  $K$  mutually  $\perp$  vectors then the ~~second term of the~~ objective can be rewritten as:

$$\sum_{k=1}^K h_k^T K h_k$$

$h_k$  can be <sup>expressed</sup> derived from the eigenvectors of  $K$ :

$$K = U \Lambda U^T \Rightarrow H = U_{[1:K]} R$$

where  $R$  is a rotation inside the eigenvalue space. ( $R^T R = I$ )

Then  $\text{tr}[H^T K H] = \sum_{k=1}^K \lambda_k$ , where  $\lambda_k$  are the largest  $K$  eigenvalues.

(- Kernel PCA is similarly solved by the  $K$  largest eigenvalues)

So after we obtain  $H$  and then  $Z = H \cdot L^{-1/2}$

We threshold the values of  $Z$   $\begin{matrix} \textcircled{1} \\ \textcircled{0} \end{matrix}$  and obtain approx. clustering solution.

It works better if we first normalize  $H$ :

$$\textcircled{H_{nk}} = \frac{H_{nk}}{\sqrt{\sum_k H_{nk}^2}}$$

THANK YOU FOR  
YOUR ATTENTION!

БЛАГОДАРЯ ЗА  
ВНИМАНИЕТО!