# W4.GMM

Phuong Anh Trinh - 11200417

February 2023

## 1 Ex 3: Compare pros and cons of Kmean, GMM

| | Kmean | GMM |
|---|---|---|
| Pros | - Relatively simple to implement<br>- Scales to large data sets<br>- Guarantees convergence<br>- Can warm-start the positions of centroids<br>- Easily adapts to new examples<br>- Generalizes to clusters of different shapes and sizes, such as elliptical clusters | - Probabilistic estimates of belonging to each cluster<br>- Does not assume spherical clusters<br>- Handles clusters of differing sizes<br>- Less sensitive to scale<br>- Handle missing data |
| Cons | - Need to choose k manually<br>- Being dependent on initial values<br>- Trouble with clustering data of varying sizes and density<br>- Sensitive to outliers<br>- Scaling with number of dimensions | - Sensitivity to initialization<br>- Assumption of normality<br>- Hard to choose number of components<br>- Struggle with high-dimensional data<br>- Limited expressive power |