

CFE for Clustering Definition

Counterfactuals for Classification

- Given
 - a classification model
 - An example y (factual) assigned to a class
- Counterfactual example (CFE) z :
 - A close example to y but assigned to different class
- Feature differences between y and z , suggest the modifications that should made to y for its class assignment to change.

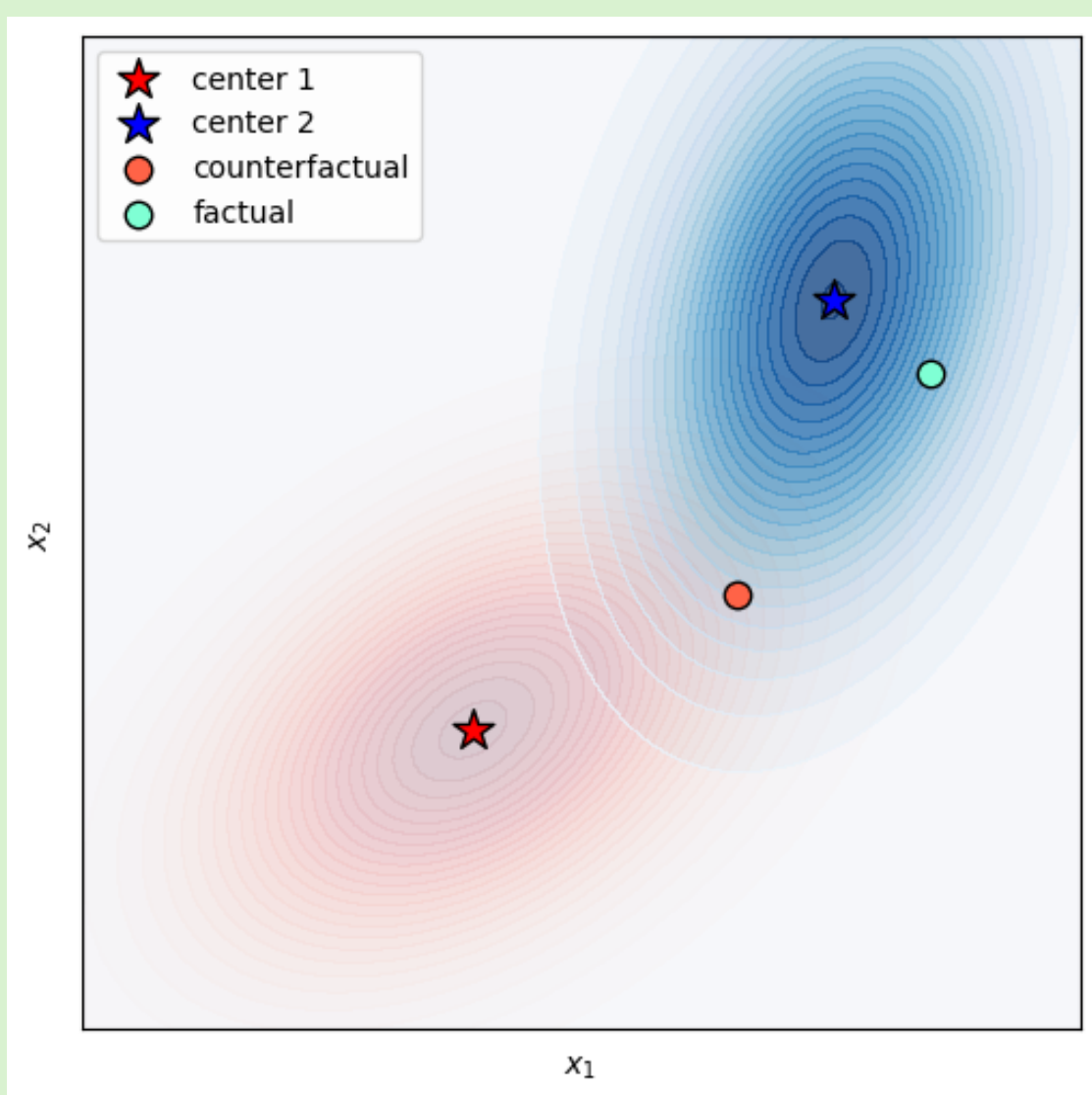
CFE for Clustering - General formulation

- A general optimization problem formulation is presented assuming:
 - A probabilistic cluster model and cluster assignment rule
 - A preference function expressing proximity and feature constraints (actionability)
 - A plausibility constraint based on cluster density

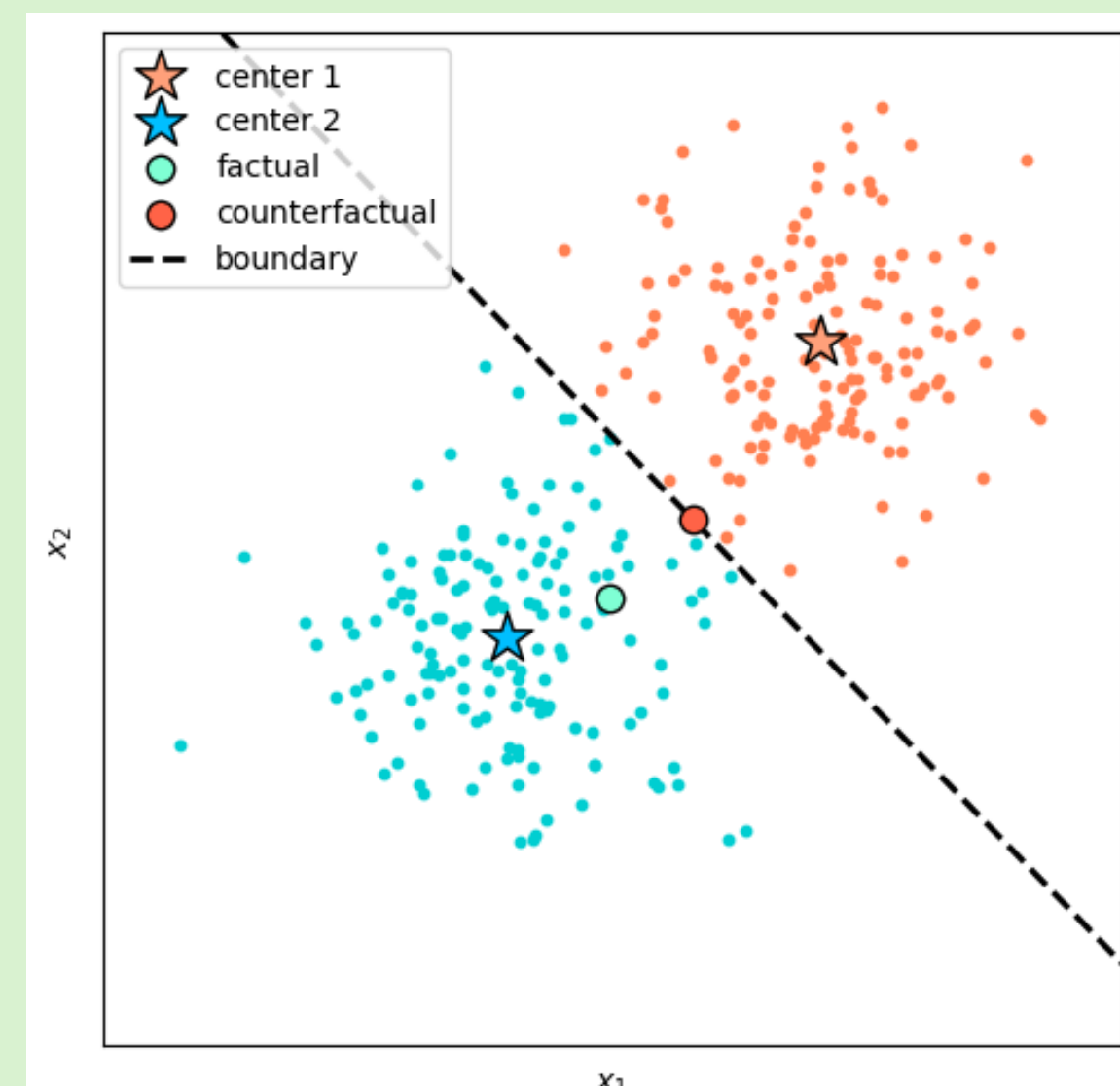
CFE for Clustering Solutions

- Assuming:
 - The CFE lies on the **cluster boundary**
 - proximity is expressed in terms squared **Euclidean distance**
- **Optimal** CFE solutions can be easily computed
 - **Analytical** (k-means model)
 - Solving a non-linear equation with **only one parameter** (Gaussian model)

CFE Examples



Gaussian Clusters

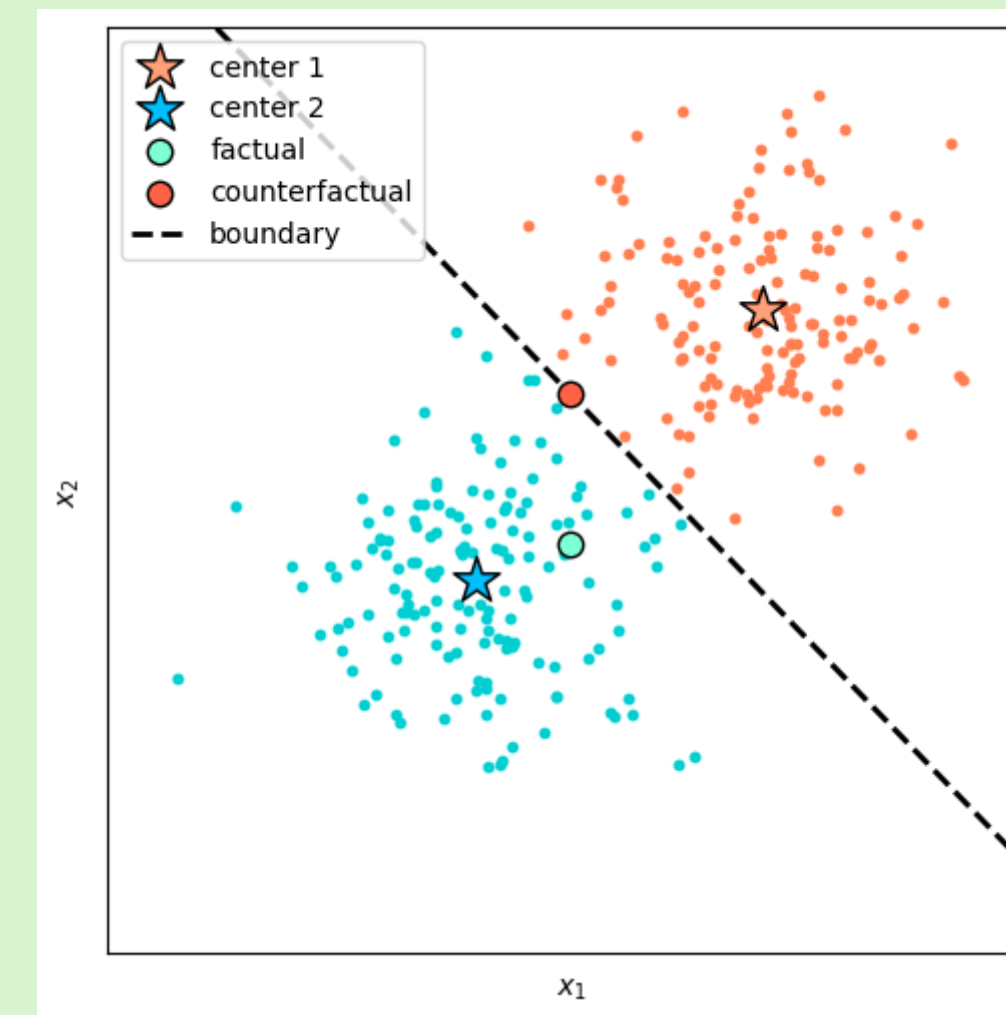
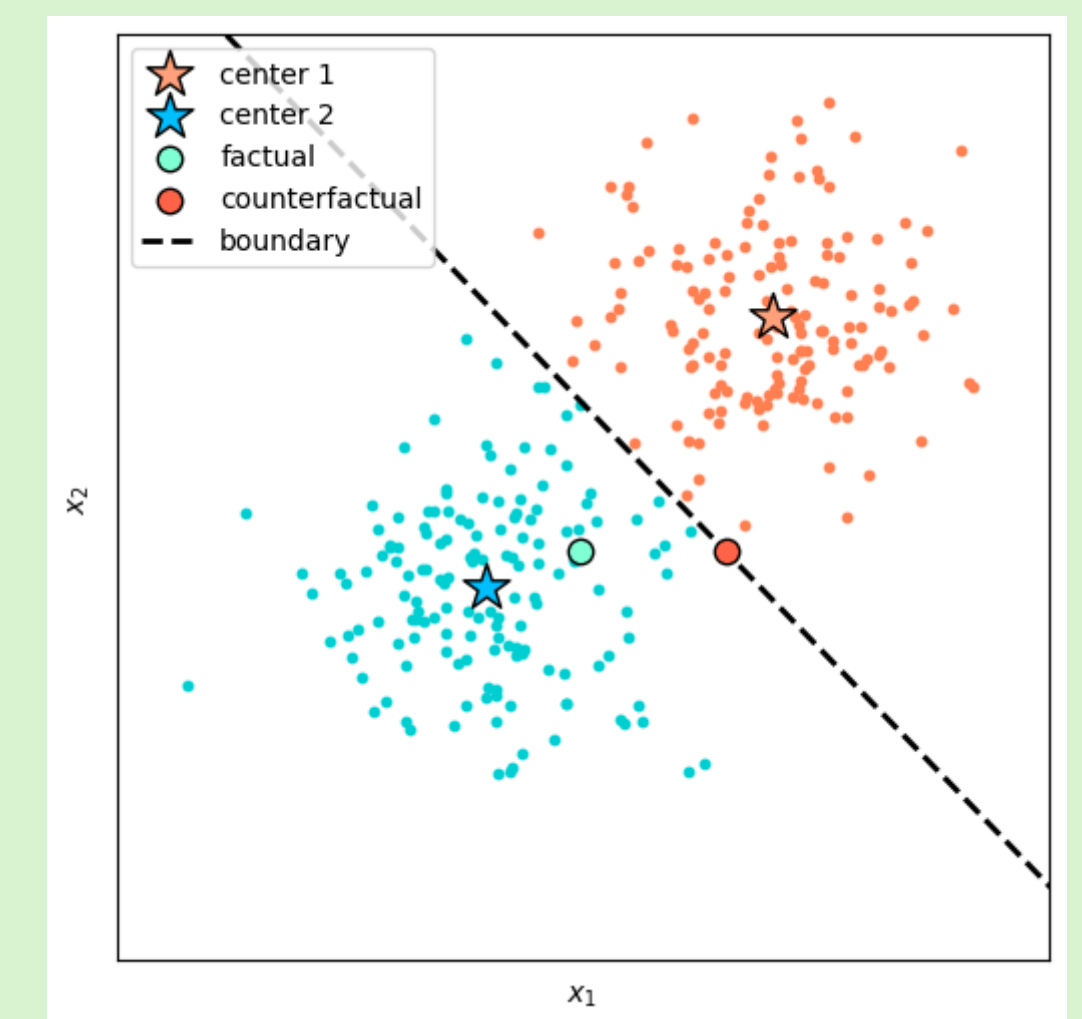


k-means

CFE solutions - Extensions

- Easily computed **optimal solutions** are also presented for:
 - Feature actionability (feature masking)
 - Increasing plausibility (moving away from the boundary towards the cluster center)
 - Mahalanobis distance for proximity (generalization of Euclidean distance)
 - CFE for group of factuals

k-means CFE Examples – Immutable features

 x_1 immutable x_2 immutable

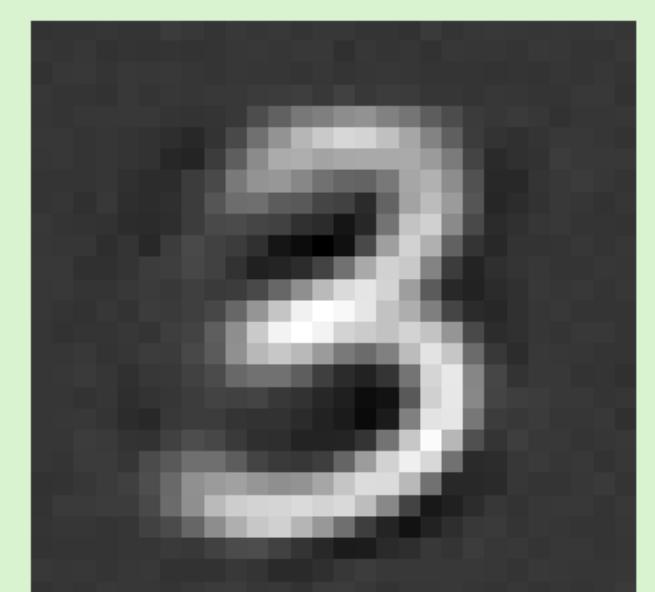
CFE for Deep Clustering

- An **encoder** neural network is used to map data instances to a **latent space**
- Clustering is performed in the latent space
- Cluster model is available in the latent space
- The factual is mapped to the latent space where the corresponding **latent-CFE** is computed
- A methodology to **project the latent-CFE back to the input space** is proposed to compute CFEs

CFE for Deep Clustering - Example



factual



CFE images of increasing plausibility

Summary

- The first approach for computing CFEs for clustering
- Fast and easy to compute optimal solutions (analytical or single parameter equations)
- Actionability and plausibility are naturally enforced
- Deep clustering extensions have been devised
- For more information:
 - <https://www.cse.uoi.gr/~fxc>