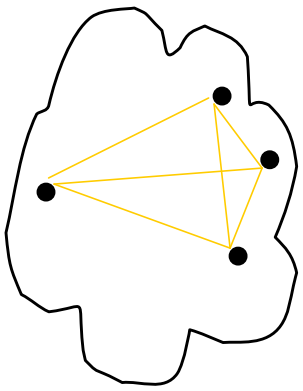
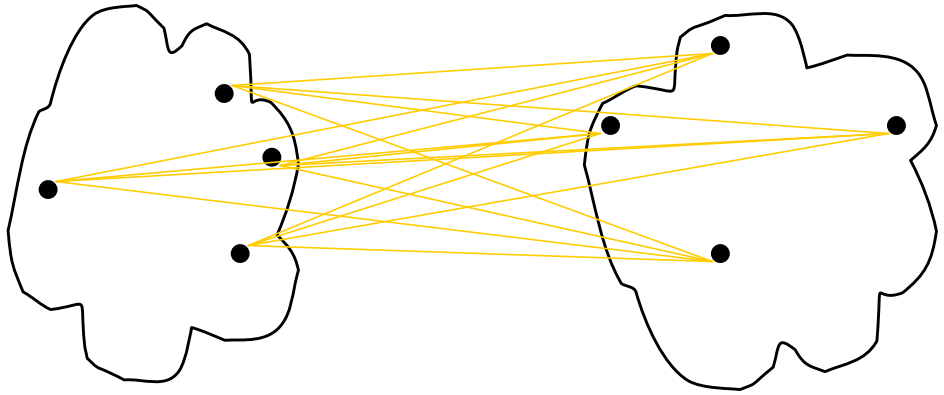
Silhouette coefficient

[Kaufman&Rousseeuw, 1990]

- **Cohesion**: measures how close objects are in a cluster
- **Separation**: measure how separated the clusters are



cohesion



separation

Silhouette coefficient

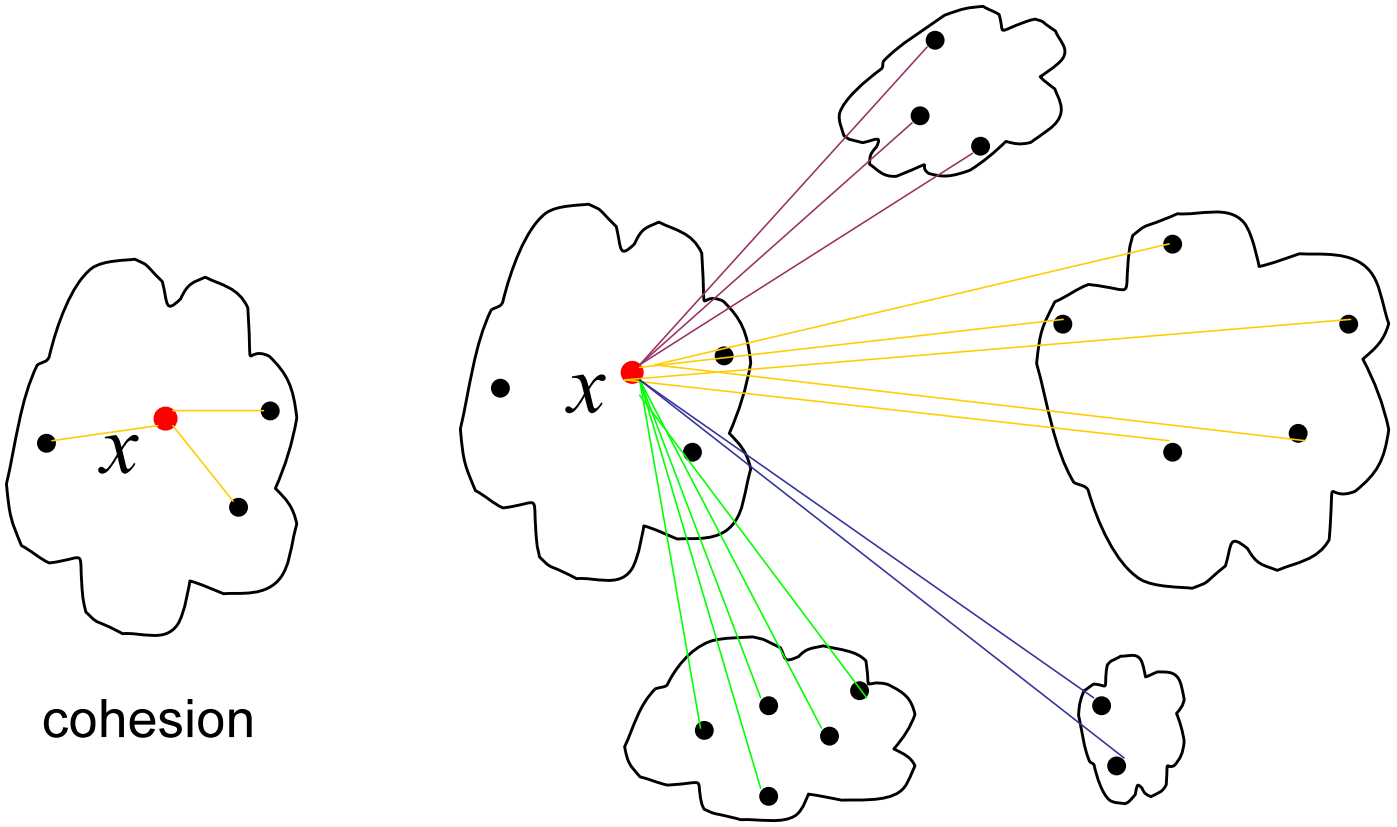
- *Cohesion* $a(x)$: average distance of x to all other vectors in the same cluster.
- *Separation* $b(x)$: average distance of x to the vectors in other clusters. Find the minimum among the clusters.
- *silhouette* $s(x)$:

$$s(x) = \frac{b(x) - a(x)}{\max\{a(x), b(x)\}}$$

- $s(x) = [-1, +1]$: -1=bad, 0=indifferent, 1=good
- Silhouette coefficient (SC):

$$SC = \frac{1}{N}\sum_{i=1}^{N} s(x)$$

Silhouette coefficient (SC)



cohesion

separation

$a(x)$: average distance in the cluster

b(x): average distances to others clusters, find minimal