Multi objective clustering

Contents

- Single objective clustering
- Combining objectives
- Multiple objectives
  - Pareto optimality
- Multi objective clustering
- LANDSAT example
- Selecting best result

Single objective

$$minimize_x \, f(x)$$

f = objective function

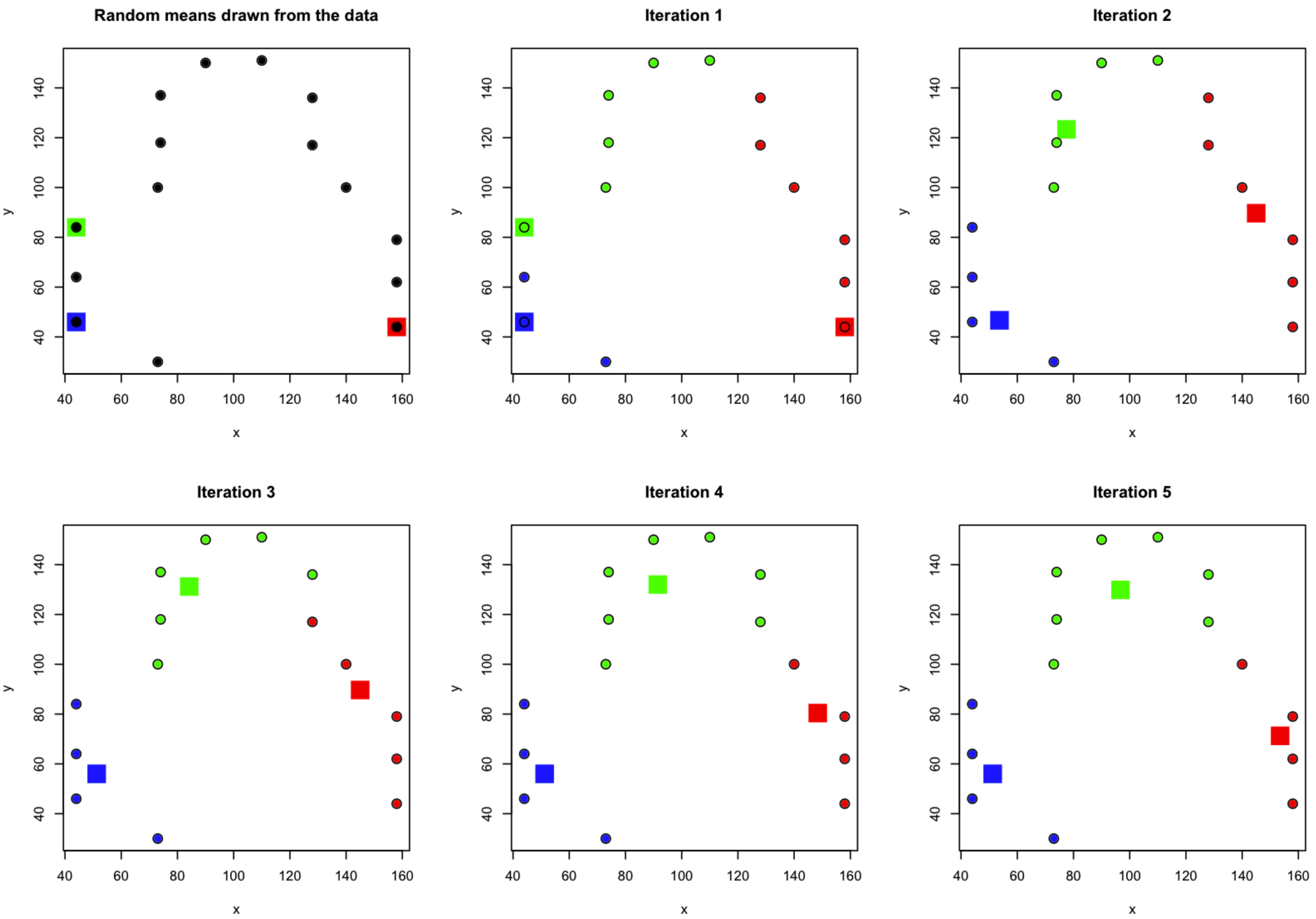
$x \in X$ = some region of n-dimensional space

EXAMPLES of f :

Cost, time, loss, …

Example: k-Means

- Set of observations $\mathrm{X} = [x_1, x_2, \ldots x_n]$
- Minimize inertia, $\sum_{i=0}^{n} \min_{\mu_j \in C}(\left\|x_i - \mu_j\right\|^2)$


Random means drawn from the data
Iteration 1
Iteration 2
Iteration 3
Iteration 4
Iteration 5
x
y
40 60 80 100 120 140 160
40 60 80 100 120 140

More than one objective

$$minimize_x \{f(x), g(x), h(x), \dots\}$$

If objectives are not conflicting, solution is trivial

EXAMPLE:

Cost vs. time

Combining objectives

- Cost and time
  - Average
  - Weighted Average
  - ε-constraint method
  - Evangelista
  - ….
- Hotel problem: Cost vs. star review?

Combining objectives

- Constructing joined objective:
  1. Model building
  2. Decision making (preference articulation)
  3. Optimization

Combining objectives

- Multi objective solution:
  1. Model building
  2. Optimization
  3. Decision making (preference articulation)

Notation

$$minimize_x \{f_1(x), f_2(x), \dots f_k(x)\}$$

- $x \in X$ are *decision vectors*
- $f(x) = (f_1(x), f_2(x), \dots f_k(x))$ are *objective vectors*
- *Optimality* of objective vector:

  no component can be improved without deteriorating another.

Notation – Pareto set

x' is *Pareto optimal* if there is no x such that $\mathrm{f}_i(x) \leq \mathrm{f_i}(\mathrm{x}')$ for all i.
Respectively it is *Weakly Pareto optimal* if the inequality is sharp.

- Set of all Pareto optimal solutions is $\mathrm{P}(\mathrm{X})$, respectively $\mathrm{WP}(\mathrm{X})$

Notation – Pareto set

x' is *Pareto optimal* if there is no x such that $\mathrm{f}_i(x) \leq \mathrm{f_i}(\mathrm{x}')$ for all i.
Respectively it is *Weakly Pareto optimal* if the inequality is sharp.

- Set of all Pareto optimal solutions is $\mathrm{P}(\mathrm{X})$, respectively $\mathrm{WP}(\mathrm{X})$
- Pareto optimal sets are *unordered*
- Pareto optimal $\sim$ *non dominated*

Notation – Pareto set

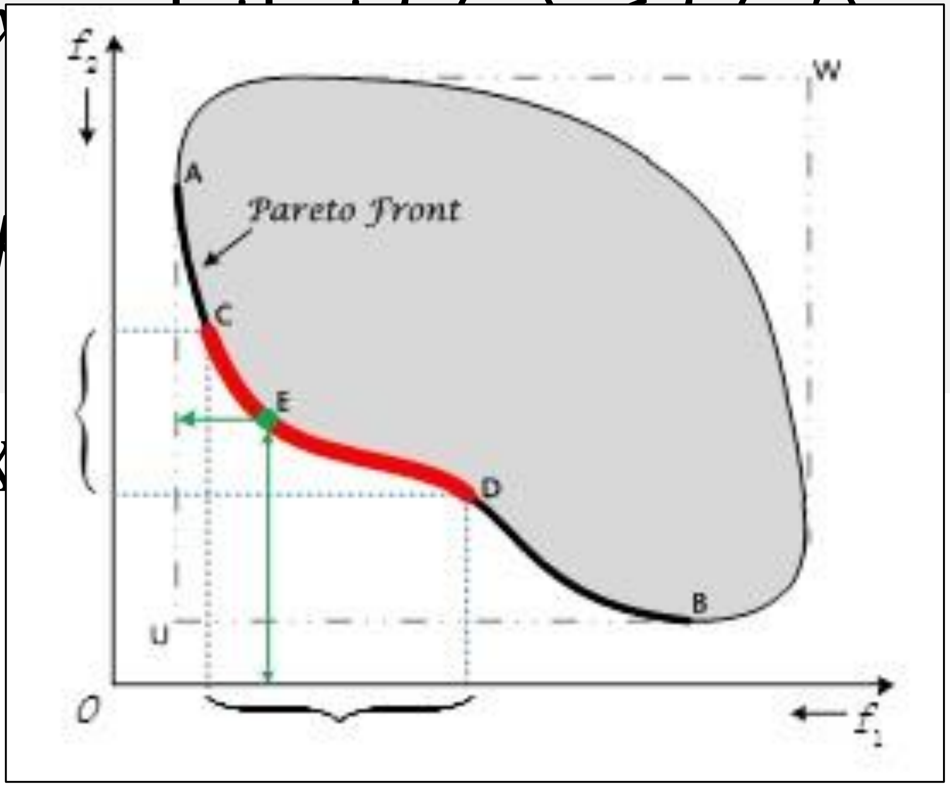
x' is *Pareto optimal* if there is no x such that $\mathrm{f}_i(x) \leq \mathrm{f_i(x')}$ for all i.
Respectively it is *Weakly Pareto optimal* if the inequality is sharp.

- Set of all Pareto optimal solutions is $\mathrm{P(X)}$, respectively $\mathrm{WP(X)}$
- Pareto optimal sets are *unordered*
- Pareto optimal $\sim$ *non dominated*

- If the space X is continuous, P(X) can contain infinite number of solutions

Notation – Pareto set

x' is *Pareto optimal* if there is no x
for all i.
Respectively it is *Weakly Pareto o*
sharp.

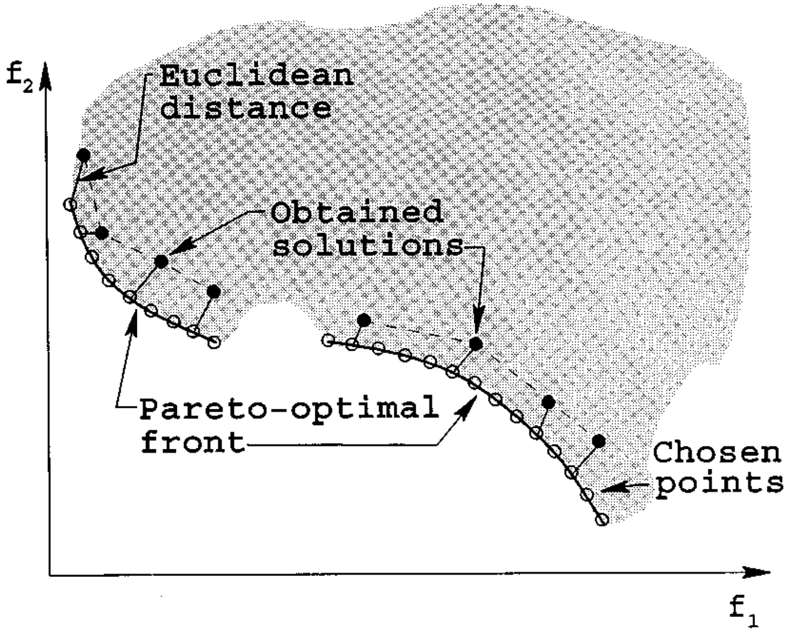


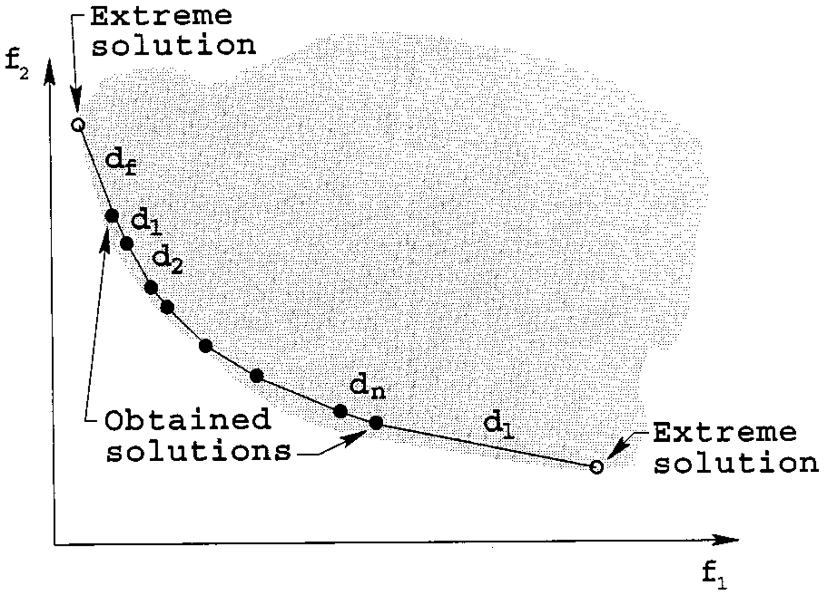
- Set of all Pareto optimal solutions is $\mathrm{P}(X$
- Pareto optimal sets are *unordered*
- Pareto optimal ~ *non dominated*

- If the space X is continuous, P(X) can contain infinite number of solutions

Multi objective clustering – NSGA-II

- Algorithm for generating Pareto front





Multi objective clustering

- Objectives:

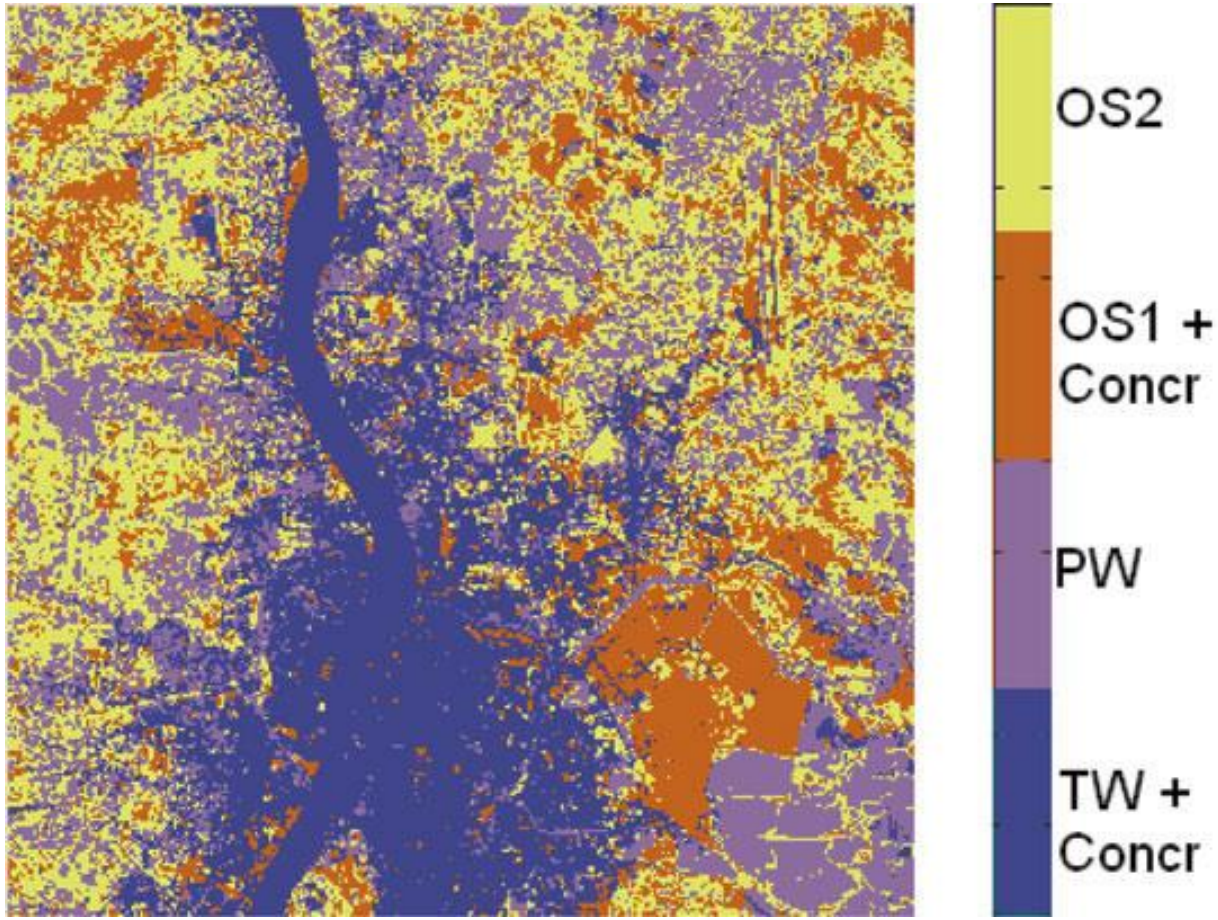
Maximize $$sep = \min_{k \neq l}\{D^2(\mu_k, \mu_l)\}$$

Minimize $$J_2 = \sum_{k=1}^{\#C} \sum_{i=1}^{n} u_{ki}^2 \, D^2(\mu_k, x_i)$$

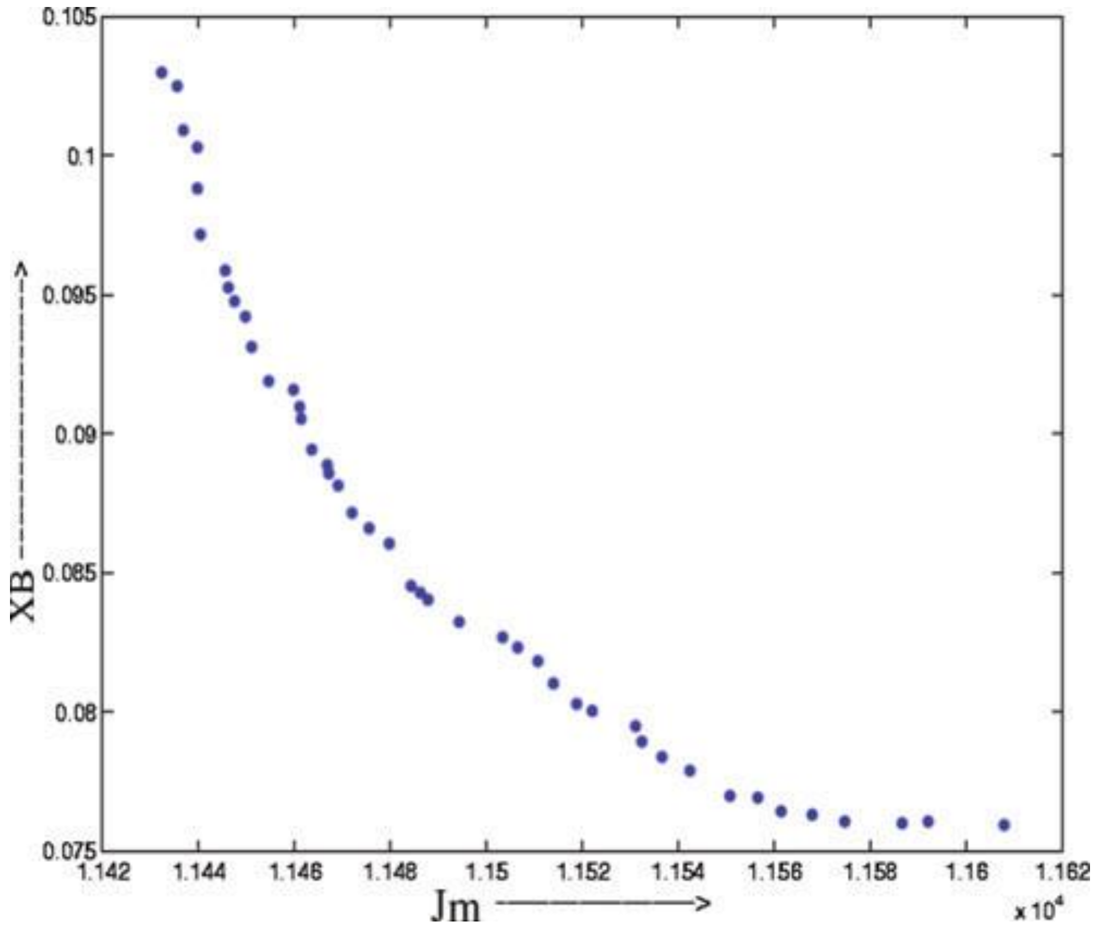
Where u_{ki} is the membership of node x_i to cluster k.

LANDSAT dataset

LANDSAT – single objective



LANDSAT – multi objective



Which solution is the best?

Problem – which solution is the best?

- Example: Cars, trains and planes

Multi objective clustering

- In case of clustering, solutions are vectors (matrices) of cluster membership.

Multi objective clustering

- In case of clustering, solutions are vectors (matrices) of cluster membership.

	x_1	x_2	x_3	x_4	x_5
Solution 1	1	1	2	3	3
Solution 2	1	2	2	3	1
Solution 3	1	3	2	3	3
Solution 4	1	1	1	3	3

Multi objective clustering

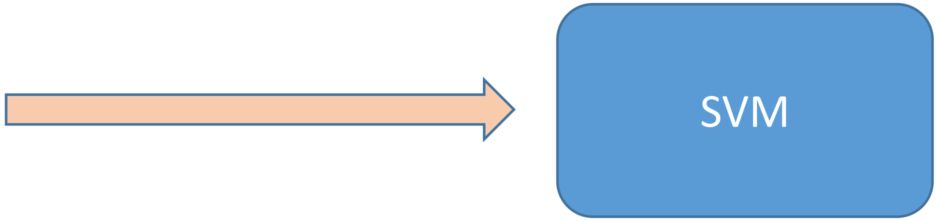
- In case of clustering, solutions are vectors (matrices) of cluster membership.

	x_1	x_2	x_3	x_4	x_5
Solution 1	1	1	2	3	3
Solution 2	1	2	2	3	1
Solution 3	1	3	2	3	3
Solution 4	1	1	1	3	3

Multi objective clustering

- In case of clustering, solutions are vectors (matrices) of cluster membership.

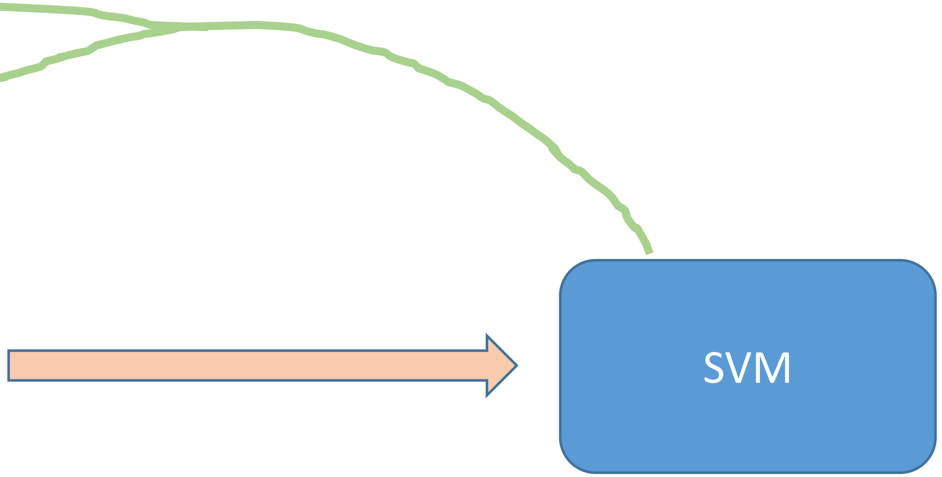
	x_1	x_2	x_3	x_4	x_5
Solution 1	1	1	2	3	3
Solution 2	1	2	2	3	1
Solution 3	1	3	2	3	3
Solution 4	1	1	1	3	3



Multi objective clustering

- In case of clustering, solutions are vectors (matrices) of cluster membership.

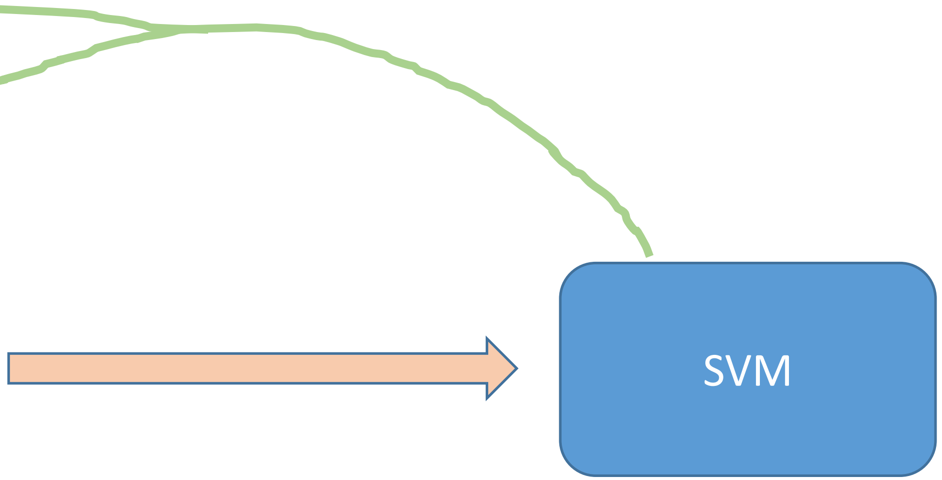
	x_1	x_2	x_3	x_4	x_5
Solution 1	1	1	2	3	3
Solution 2	1	2	2	3	1
Solution 3	1	3	2	3	3
Solution 4	1	1	1	3	3



Multi objective clustering

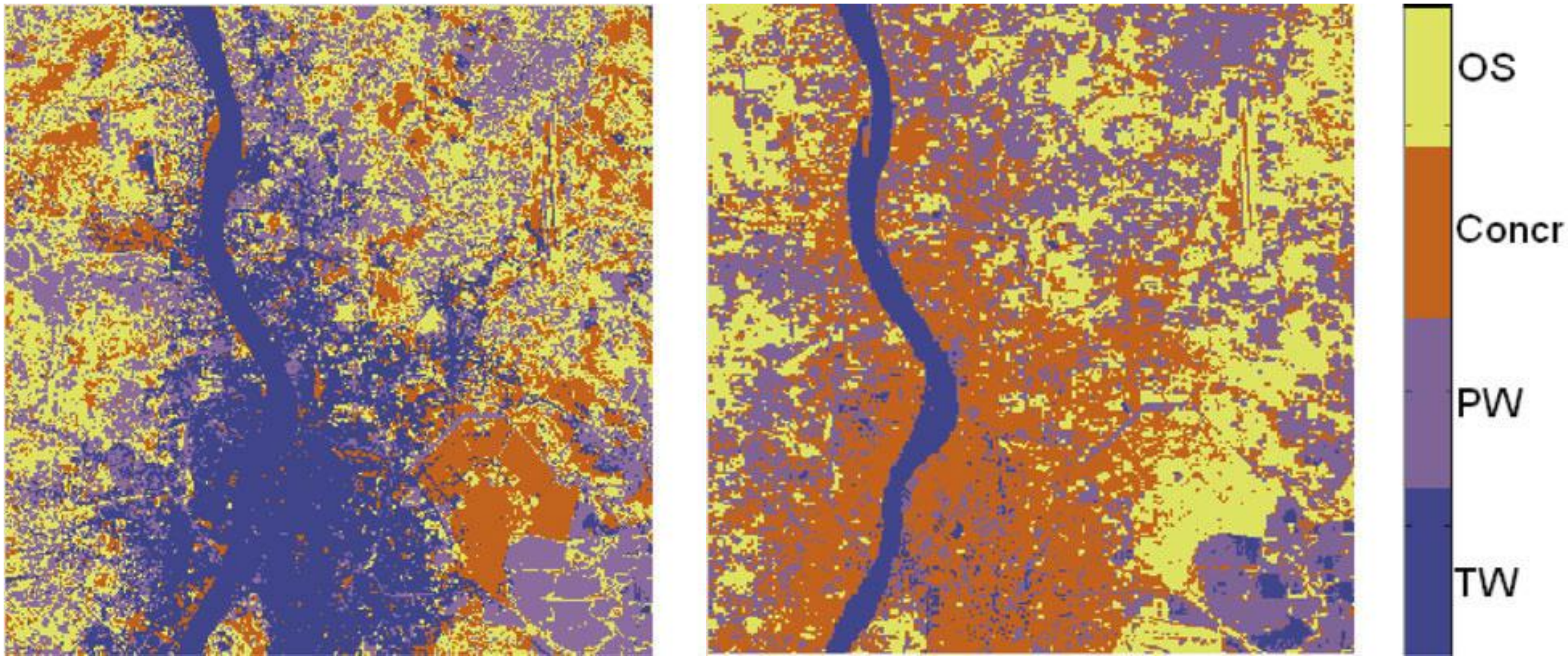
- In case of clustering, solutions are vectors (matrices) of cluster membership.

	x_1	x_2	x_3	x_4	x_5
Solution 1	1	1	2	3	3
Solution 2	1	2	2	3	1
Solution 3	1	3	2	3	3
Solution 4	1	1	1	3	3



Final solution	**1**	**1**	**2**	**3**	**3**

LANDSAT – multi objective



Conclusion

- Multiple objectives can perform better than single objective
- Selecting objectives is important

References

- K. Deb, A. Pratap, S. Agarwal and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," in *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182-197, Apr 2002.
- Branke, Jürgen, et al., eds. *Multiobjective optimization: Interactive and evolutionary approaches*. Vol. 5252. Springer, 2008.
- Maulik, Ujjwal, Sanghamitra Bandyopadhyay, and Anirban Mukhopadhyay.*Multiobjective Genetic Algorithms for Clustering: Applications in Data Mining and Bioinformatics*. Springer Science & Business Media, 2011.

Fuzzy C-Means

- Initialize U=[uij] matrix, U(0)
- At k-step: calculate the centers vectors C(k)=[cj] with U(k)

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^{m} \cdot x_i}{\sum_{i=1}^{N} u_{ij}^{m}}$$

- Update U(k) , U(k+1)

$$u_{ij} = \frac{1}{\sum_{k=1}^{C} \left(\frac{\| x_i - c_j \|}{\| x_i - c_k \|} \right)^{\frac{2}{m-1}}}$$

- If || U(k+1) - U(k)||< epsilon then STOP; otherwise return to step 2.