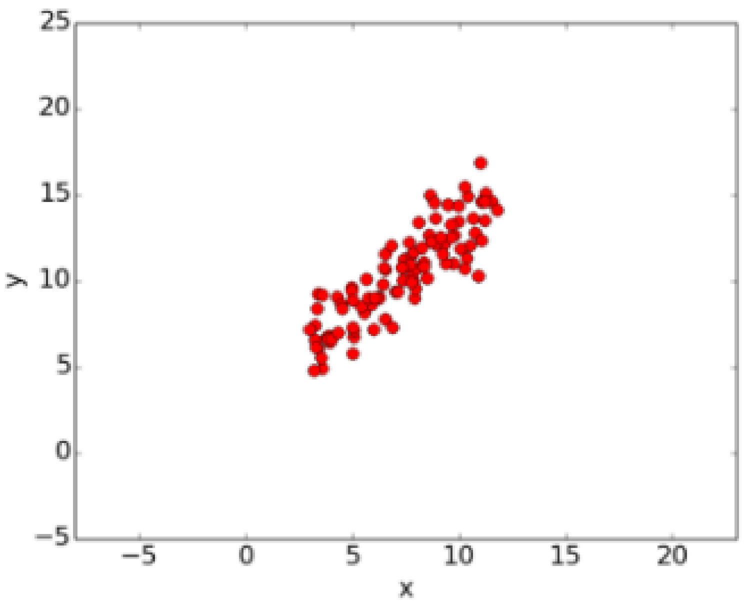
LOCAL SEARCH ALGORITHMS: Applications

Regression Problem, K-Means Clustering Problem.

Linear regression – an example that uses gradient descent



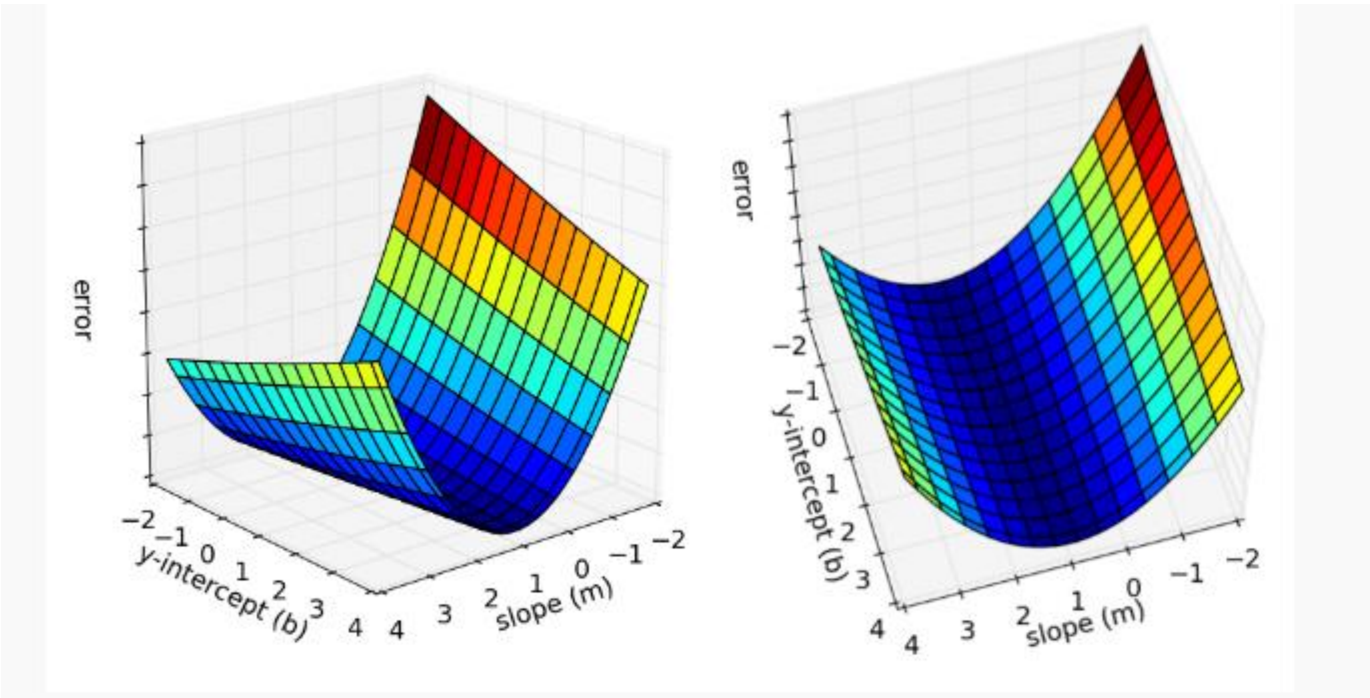
$$y = mx + b$$

We want to fit a straight line (in 2D case). The sample we are given with is $\{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$.

We want to find (m, b).

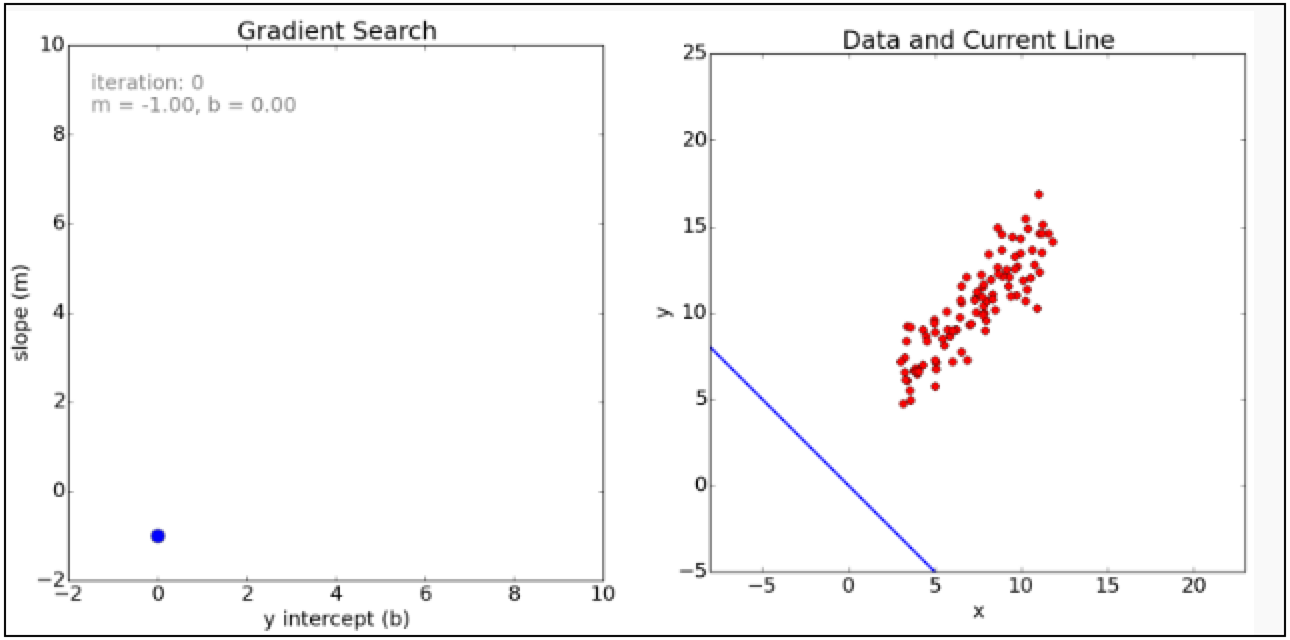
In the space (m, b) what is the function we are going to define? Minimum value of that function should be a solution for us.

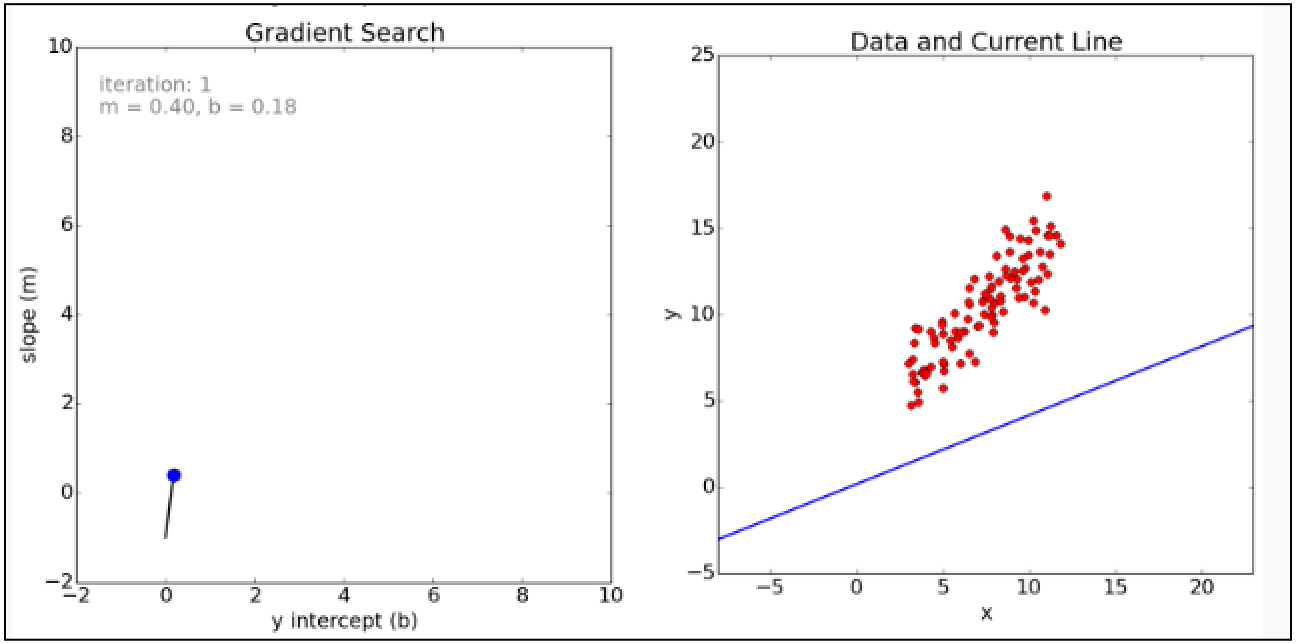
$$\text{Error}_{(\text{m,b})} = \frac{1}{N}\sum_{i-1}^{N}(y_i - (mx_i + b))^2$$

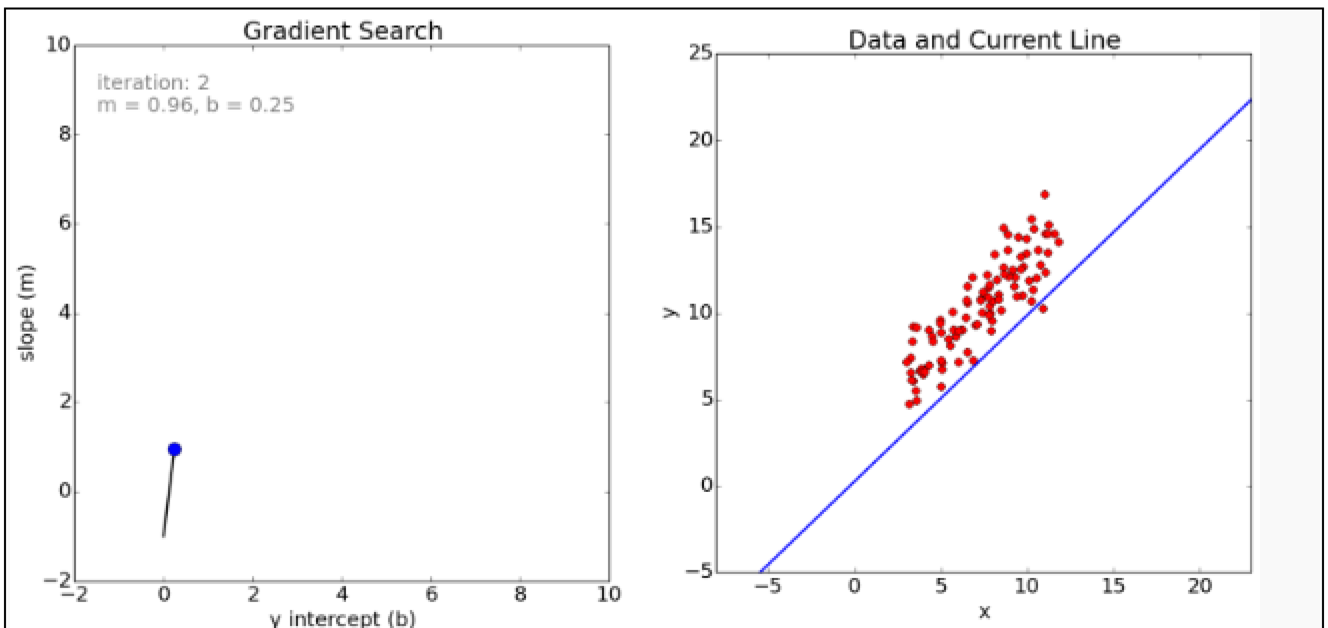


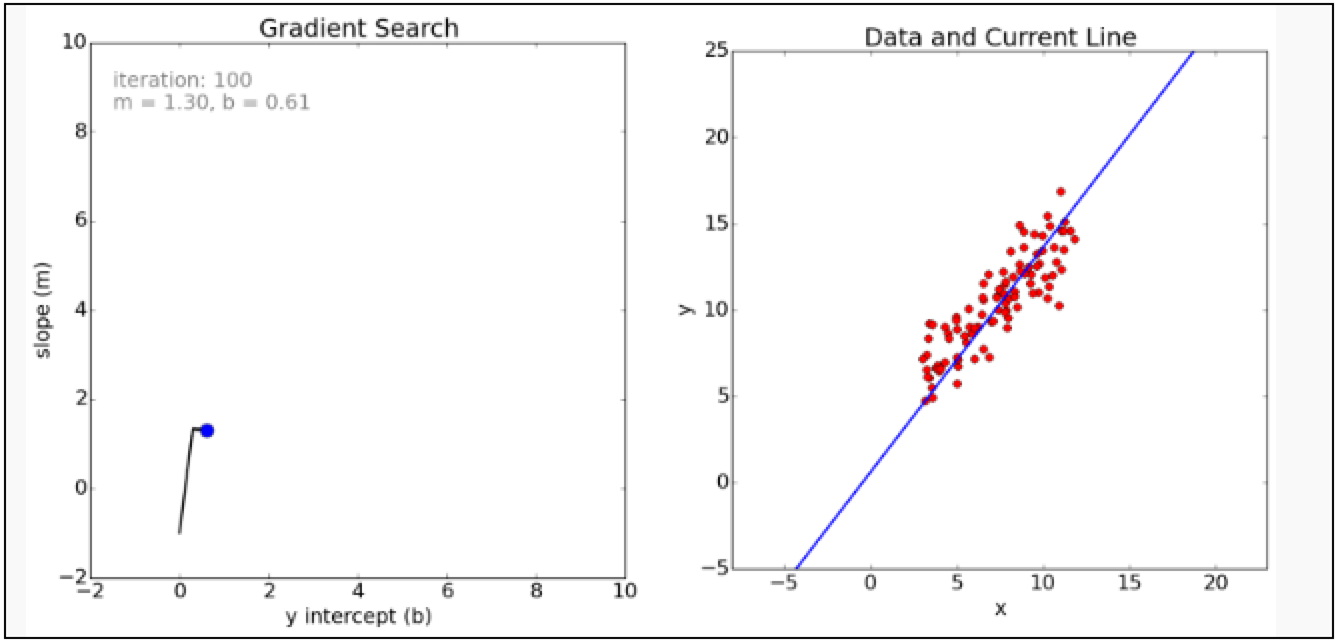
$$\frac{\partial}{\partial \mathrm{m}} = \frac{2}{N}\sum_{i=1}^{N} -x_i(y_i - (mx_i + b))$$

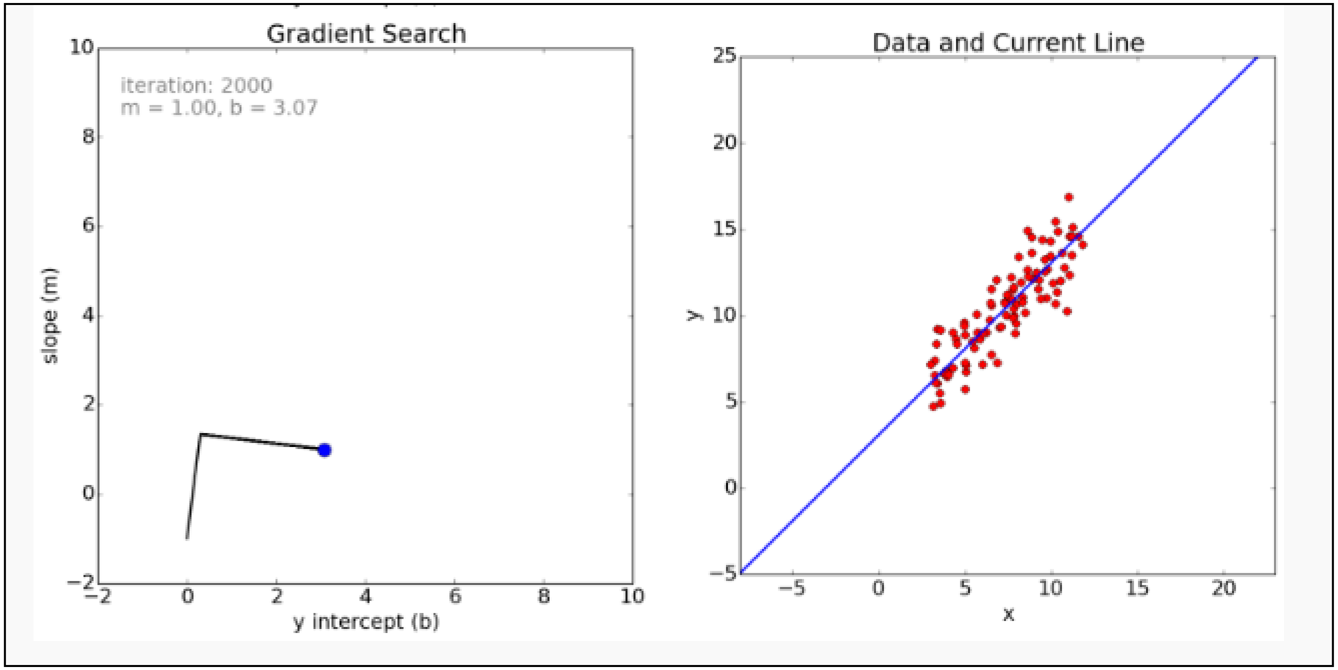
$$\frac{\partial}{\partial \mathrm{b}} = \frac{2}{N}\sum_{i=1}^{N} -(y_i - (mx_i + b))$$




Gradient Search
iteration: 1
m = 0.40, b = 0.18
slope (m)
y intercept (b)
Data and Current Line
y
x


Gradient Search
iteration: 2
m = 0.96, b = 0.25
slope (m)
y intercept (b)
Data and Current Line
y
x


Gradient Search
iteration: 100
m = 1.30, b = 0.61
slope (m)
y intercept (b)
Data and Current Line
y
x

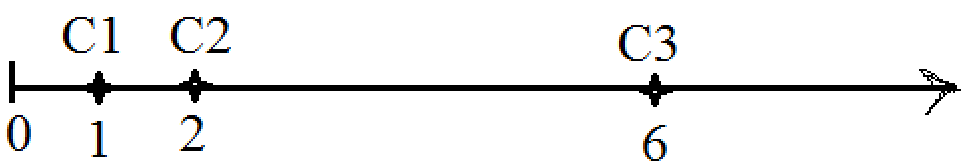

Gradient Search
iteration: 2000
m = 1.00, b = 3.07
slope (m)
y intercept (b)
Data and Current Line
y
x

Closed form solution

- Since the criterion that is minimized is quadratic, the linear regression problem must be having a closed form solution.
- This is nothing but applying the Newton's descent method.

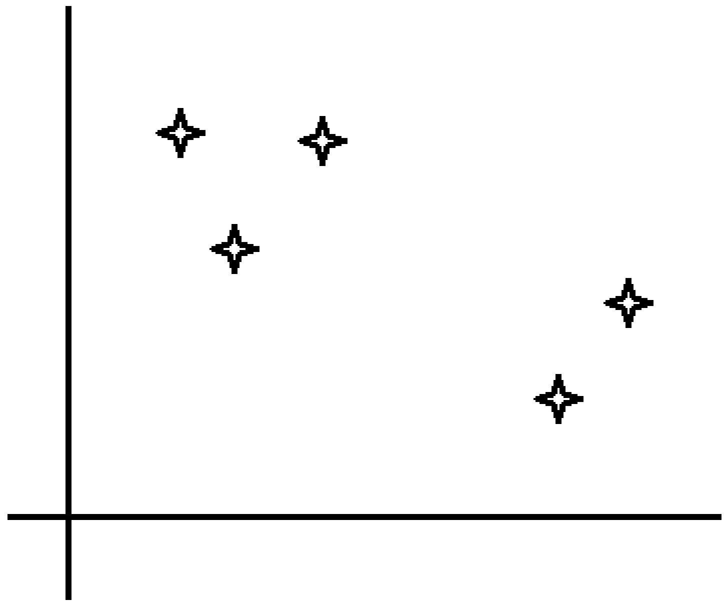
Single airport problem in 1D

- We are asked to locate the airport place, so that it is "sum of squared distances" from all 3 cities, viz., C1,C2, and C3, is minimized.



2 dimensional, single airport

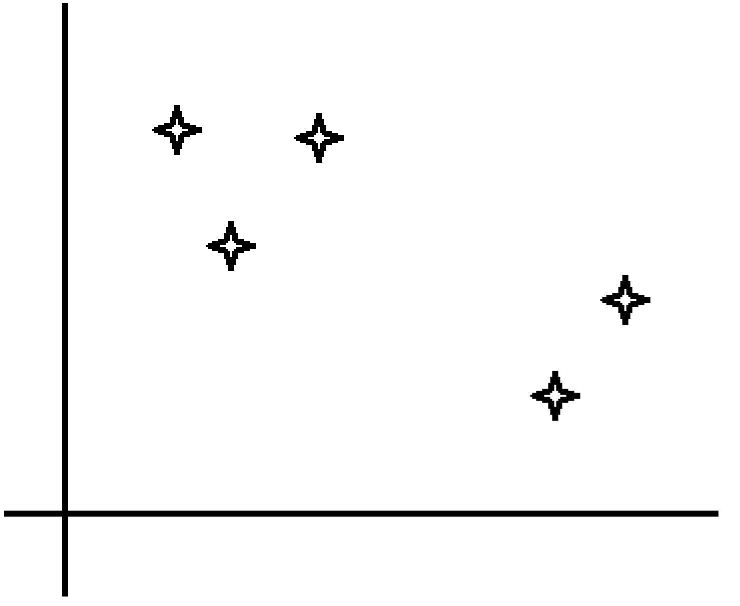
- Closed form solution
  - The centroid?
  - Why?



2 dimensional, 2 airports problem

- **Single airport problem can be easily visualized.**

- **But 2 airports problem is difficult to visualize.**
  - **You need to work in a 4 dimensional space.**



- **3 airports problem you need to work in a 6D space.**

Airports Problem (3 airports) …

- Coordinates of the three airports be

$$(x_1, y_1)^t, (x_2, y_2)^t, (x_3, y_3)^t$$

- $f((x_1, y_1, x_2, y_2, x_3, y_3)^t)$ = Sum of squared distances from each city to its nearest airport
- Find values for the six parameters that minimize $f(\cdot)$

Note, we use column vectors to represent points in higher dimensional space

The 3 airports problem

- What is the criterion?
- Sum of squared distances of cities from their nearest airport locations.

K means clustering : Introduction

Let the cities be located at $C_1, C_2, \ldots, C_n$ and let the three airport locations be A_1, A_2, A_3 .

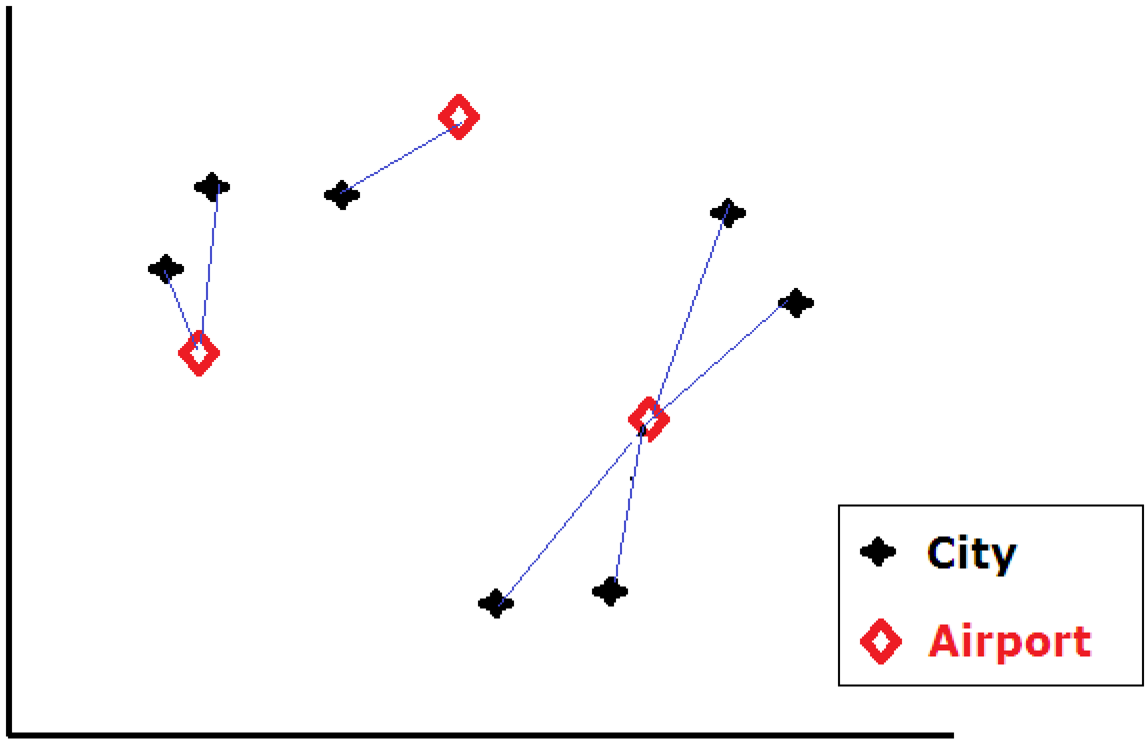
Arbitrarily choose A_1, A_2, A_3.

Let $J = \sum_i \min_j \left\|C_i - A_j\right\|^2$

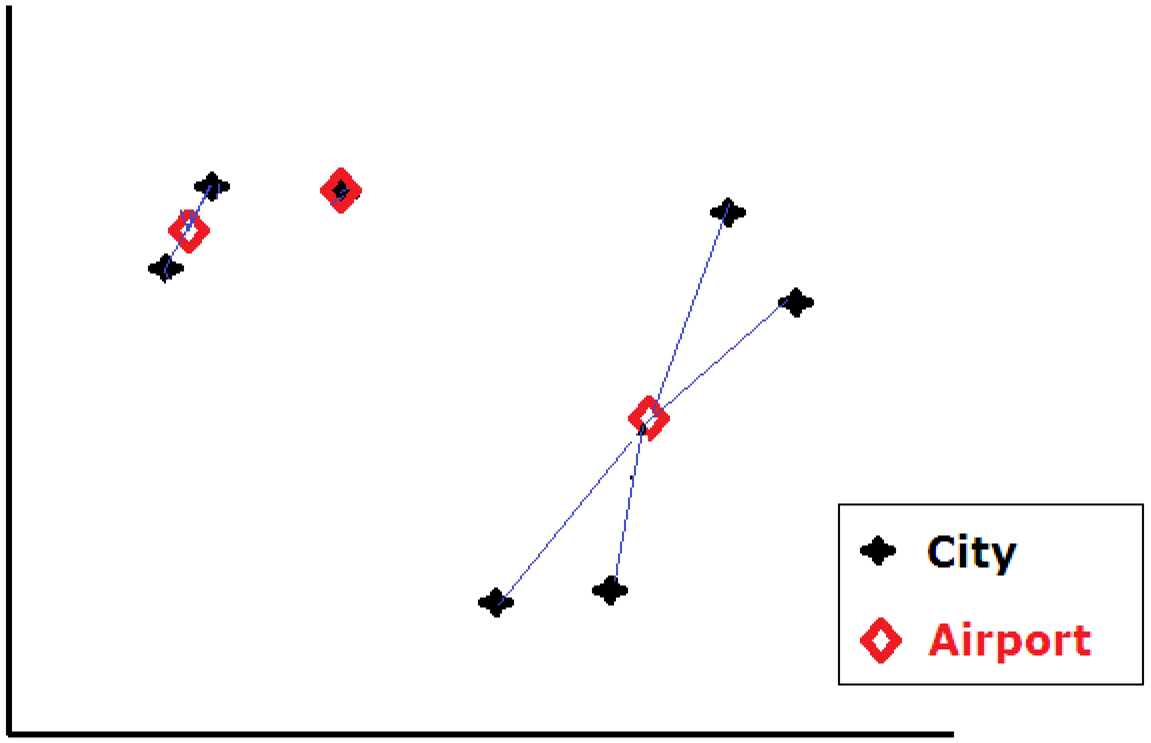
We want to find A_1, A_2, A_3 such that J is minimized.

- Minimum of convex functions, in general is not convex.
- So the objective is not convex.

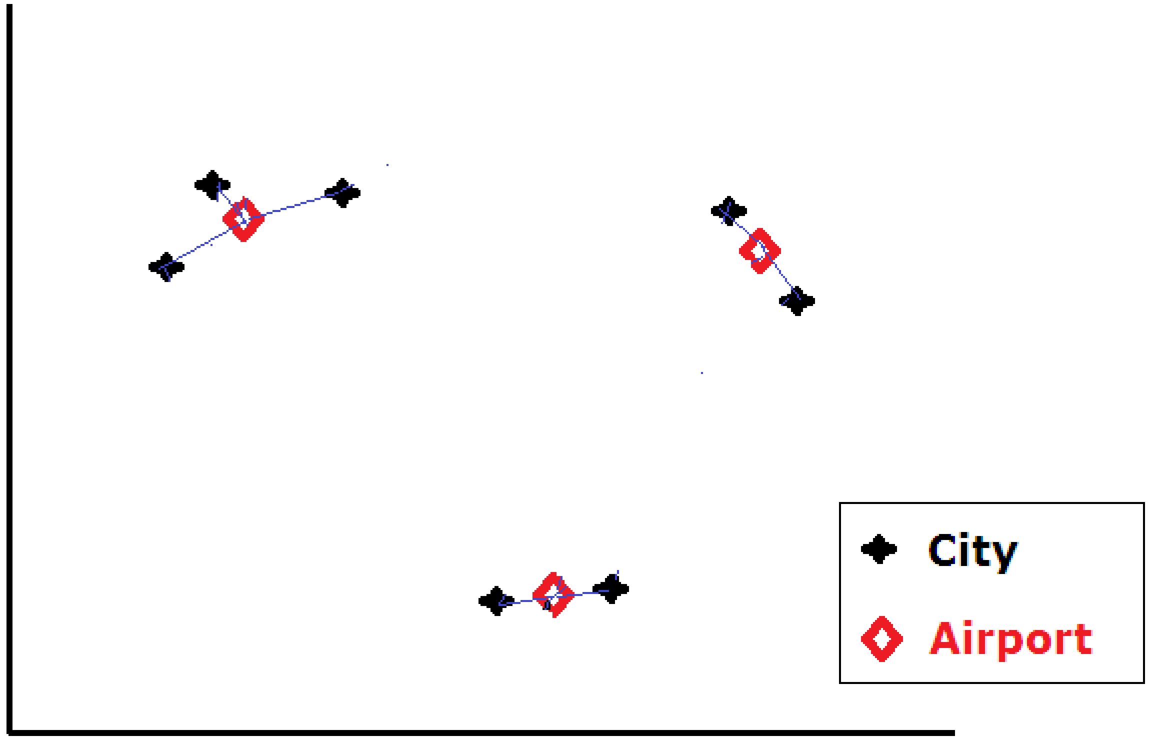
An illustration: how k means works



This, indeed, gets stuck with local minima.



Global minima is …



Lloyd in 1982 gave the k means clustering algorithm

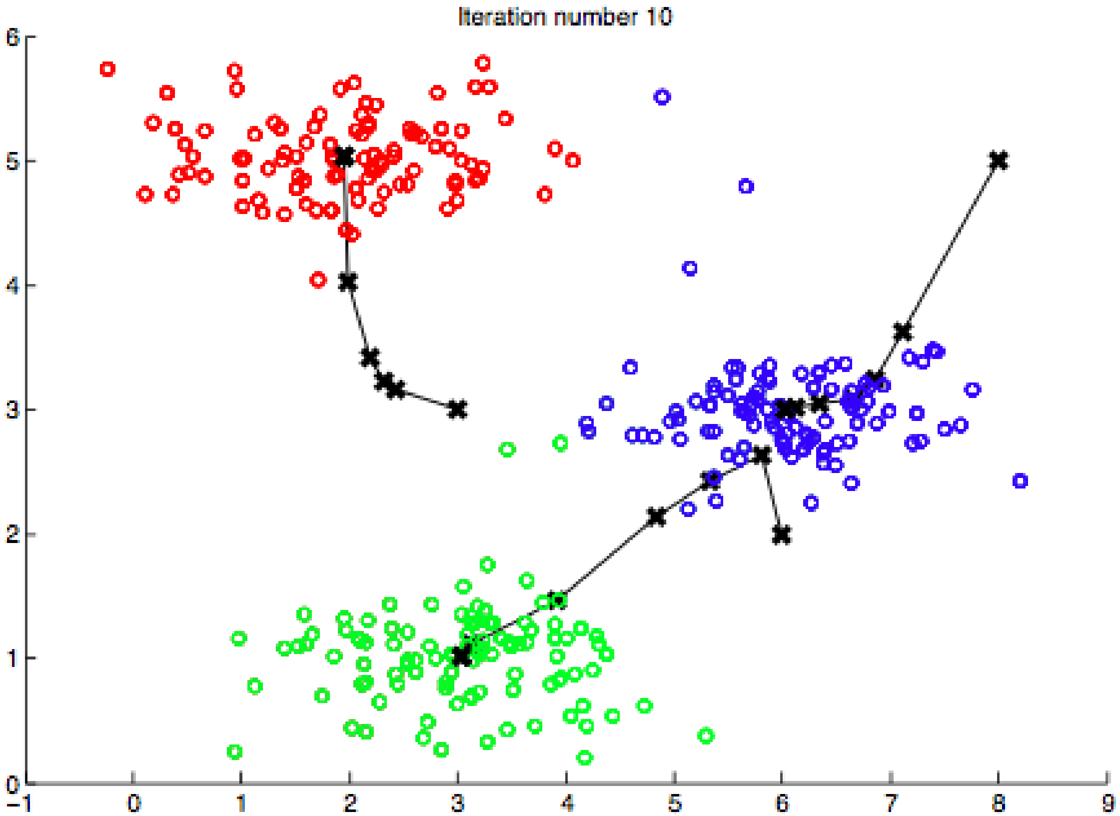
- This is an iterative Newton’s Descent Method.
- Gradient descent also works, but is slow.

- Single airport problem can be solved with a closed form solution. (Newton’s method gives this).

K airports problems via k means clustering

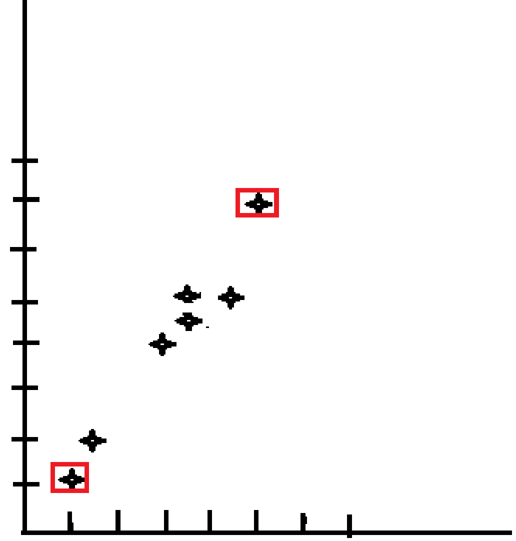
- Choose randomly k points from the data. (these are the initial airport locations)
- **(1)** Assign each point to its nearest airport (center/mean) ➔This gives partition of the data.
- **(2)** For each block of the partition, solve single airport problem. ➔Reduce the criterion using Newton’s method. (this gives k new points)
- Repeat **(1)** and **(2)** iteratively till convergence.

Illustration: see how mean vectors are moving as iterations are increased.



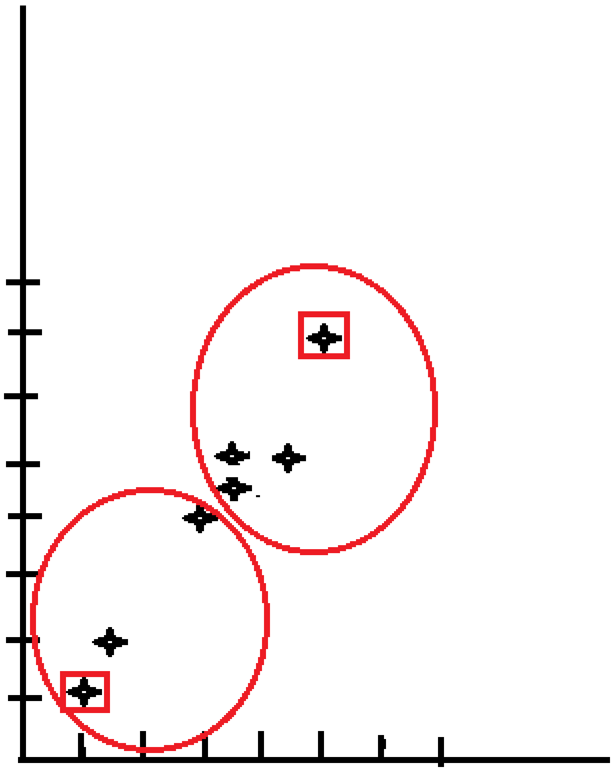
Example, k=2

Initial means :
Point 1 and Point 4



Point	x	y
1	**1.0**	**1.0**
2	1.5	2.0
3	3.0	4.0
4	**5.0**	**7.0**
5	3.5	5.0
6	4.5	5.0
7	3.5	4.5

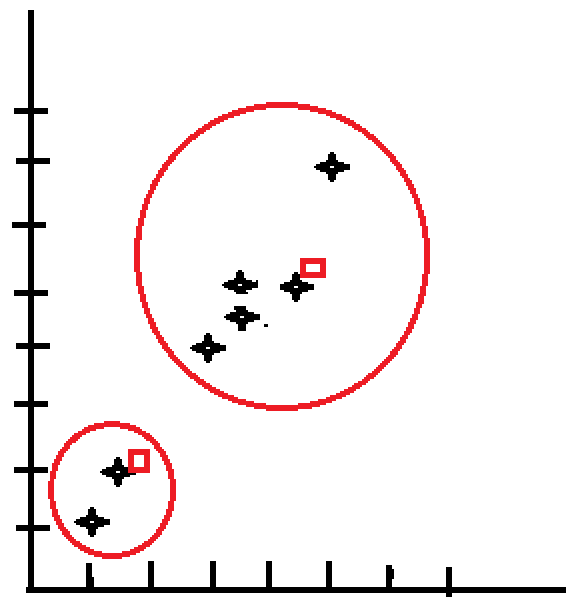
Individual	Distance to mean 1	Distance to mean 2
1	0	7.21
2	1.12	6.10
3	3.61	3.61
4	7.21	0
5	4.71	2.5
6	5.31	2.06
7	4.3	2.91

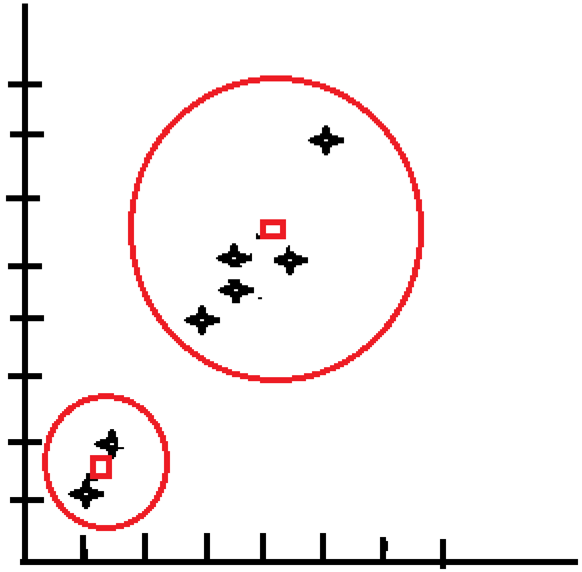


Clusters: {1,2,3}, {4,5,6,7}

New Mean vectors are:
(1.83,2.33), (4.125, 5.375)

Point 3 now is closer to
mean 2
Dist (3, mean1) = 2.039
Dist (3, mean2) = 1.777
So, point 3 moves from cluster 1 to cluster 2.
Clusters = {1,2}, {3,4,5,6,7}





Point	x	y
1	1.0	1.0
2	1.5	2.0
3	3.0	4.0
4	5.0	7.0
5	3.5	5.0
6	4.5	5.0
7	3.5	4.5

- Clusters = {1,2}, {3,4,5,6,7}
- Final means = {(1.25, 1.5),(3.9, 5.1)}