

Pick K random cluster centers,  $\mu_1 \dots \mu_k$

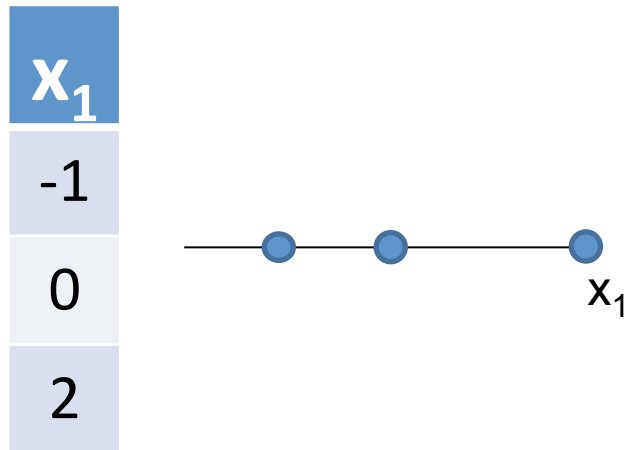
For  $t=1..T$ :

- E step:

$$p(y = i | x^j; \theta_t) \propto \exp \left( -\frac{1}{2\sigma^2} (x^j - \mu_i)^2 \right)$$

- M step:

$$\mu_i = \frac{\sum_{j=1}^m p(y = i | x^j; \theta_t) x^j}{\sum_{j=1}^m p(y = i | x^j; \theta_t)}$$



Initialization, random means and  $\sigma=1$ :

- $\mu_1 = -1, \mu_2 = 0$

$t=0$ :

- $P(y=1 | x^1) \propto \exp(-0.5 \times (-1+1)^2) = 1$
- $P(y=2 | x^1) \propto \exp(-0.5 \times (-1-0)^2) = 0.6$ 
  - $P(y=1 | x^1) = 0.63, P(y=2 | x^1) = 0.37$
- $P(y=1 | x^2) \propto \exp(-0.5 \times (0+1)^2) = 0.6$
- $P(y=2 | x^2) \propto \exp(-0.5 \times (0-0)^2) = 1$ 
  - $P(y=1 | x^2) = 0.37, P(y=2 | x^2) = 0.63$
- $P(y=1 | x^3) \propto \exp(-0.5 \times (2+1)^2) = 0.07$
- $P(y=2 | x^3) \propto \exp(-0.5 \times (2-0)^2) = 0.93$ 
  - $P(y=1 | x^3) = 0.01, P(y=2 | x^3) = 0.93$
- $\mu^1 = (0.63 \times -1 + 0.37 \times 0 + 0.07 \times 2) / (0.63 + 0.37 + 0.07) = -0.45$
- $\mu^2 = (0.37 \times -1 + 0.67 \times 0 + 0.93 \times 2) / (0.37 + 0.67 + 0.93) = 0.75$

$t=1$ :

- learning continues, when do we stop?