



Bayesian Non-Parametrics

Advanced Machine Learning for NLP

Jordan Boyd-Graber

GIBBS EXAMPLE

- Don't know how many clusters there are
- Gibbs sampling: change the assignment of one cluster conditioned on all other clusters
- Convergence harder to detect
- Course is focused on variational inference, but Gibbs more intuitive
- Equation

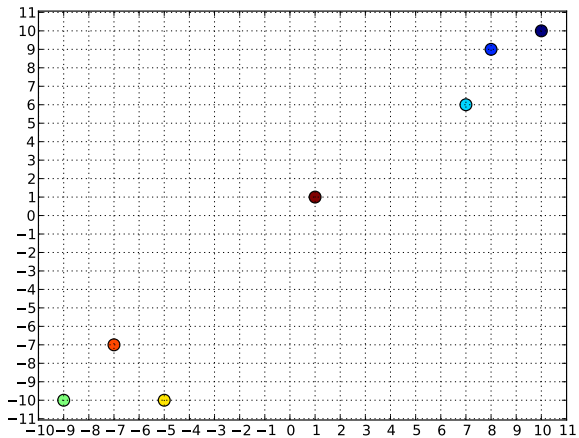
$$p(z_i = k | \vec{z}_{-i}, \vec{x}, \{\theta_k\}, \alpha) \propto \begin{cases} \left(\frac{n_k}{n+\alpha}\right) \mathcal{N}\left(x, \frac{n\bar{x}}{n+1}, 1\right) & \text{existing} \\ \frac{\alpha}{n+\alpha} \mathcal{N}(x, 0, 1) & \text{new} \end{cases} \quad (1)$$

Simplification

We'll assume that:

$$p(x | \bar{x}) \propto \exp \left\{ -\sqrt{\left(x_1 - \frac{n}{n+1} \bar{x}_1\right)^2 + \left(x_2 - \frac{n}{n+1} \bar{x}_2\right)^2} \right\} \quad (2)$$

Data



Sampling point 0

Compute the (proportional) probability of assigning data 0 to a new cluster and cluster 1.

Recall that $\alpha = 0.25$ and

$$p(x | \bar{x}) \propto \exp \left\{ -\sqrt{\left(x_1 - \frac{n}{n+1} \bar{x}_1\right)^2 + \left(x_2 - \frac{n}{n+1} \bar{x}_2\right)^2} \right\} \quad (3)$$

i	x_1	x_2	z_i
0	10	10	
1	8	9	1
2	7	6	2
3	-9	-10	3
4	-5	-10	4
5	-7	-7	5
6	1	1	6

Sampling point 0

- There are currently 6 clusters

(3)

Sampling point 0

- There are currently 6 clusters

$$p(z_0 = \text{new} | \vec{z}_{-0}, \vec{x}, \alpha) \propto \frac{0.25}{6 + 0.25} \mathcal{N} \left(\begin{matrix} 10.00 \\ 10.00 \end{matrix} \mid \begin{matrix} 0.00 \\ 0.00 \end{matrix}, \mathbf{1} \right) = 0.04 \times 0.00000 \quad (3)$$

(4)

Sampling point 0

- There are currently 6 clusters

$$p(z_0 = \text{new} | \vec{z}_{-0}, \vec{x}, \alpha) \propto \frac{0.25}{6+0.25} \mathcal{N} \left(\begin{matrix} 10.00 \\ 10.00 \end{matrix} \mid \begin{matrix} 0.00 \\ 0.00 \end{matrix}, \mathbb{1} \right) = 0.04 \times 0.00000 \quad (3)$$

$$p(z_0 = 1 | \vec{z}_{-0}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{matrix} 10.00 \\ 10.00 \end{matrix} \mid \begin{matrix} 4.00 \\ 4.50 \end{matrix}, \mathbb{1} \right) = 0.16 \times 0.00029 \quad (4)$$

(5)

Sampling point 0

- There are currently 6 clusters

$$p(z_0 = \text{new} | \vec{z}_{-0}, \vec{x}, \alpha) \propto \frac{0.25}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} 10.00 & 0.00 \\ 10.00 & 0.00 \end{array}, \mathbf{1} \right) = 0.04 \times 0.000000 \quad (3)$$

$$p(z_0 = 1 | \vec{z}_{-0}, \vec{x}, \alpha) \propto \frac{1.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} 10.00 & 4.00 \\ 10.00 & 4.50 \end{array}, \mathbf{1} \right) = 0.16 \times 0.00029 \quad (4)$$

$$p(z_0 = 2 | \vec{z}_{-0}, \vec{x}, \alpha) \propto \frac{1.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} 10.00 & 3.50 \\ 10.00 & 3.00 \end{array}, \mathbf{1} \right) = 0.16 \times 0.00007 \quad (5)$$

$$p(z_0 = 3 | \vec{z}_{-0}, \vec{x}, \alpha) \propto \frac{1.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} 10.00 & -4.50 \\ 10.00 & -5.00 \end{array}, \mathbf{1} \right) = 0.16 \times 0.000000 \quad (6)$$

$$p(z_0 = 4 | \vec{z}_{-0}, \vec{x}, \alpha) \propto \frac{1.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} 10.00 & -2.50 \\ 10.00 & -5.00 \end{array}, \mathbf{1} \right) = 0.16 \times 0.000000 \quad (7)$$

$$p(z_0 = 5 | \vec{z}_{-0}, \vec{x}, \alpha) \propto \frac{1.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} 10.00 & -3.50 \\ 10.00 & -3.50 \end{array}, \mathbf{1} \right) = 0.16 \times 0.000000 \quad (8)$$

$$p(z_0 = 6 | \vec{z}_{-0}, \vec{x}, \alpha) \propto \frac{1.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} 10.00 & 0.50 \\ 10.00 & 0.50 \end{array}, \mathbf{1} \right) = 0.16 \times 0.000000 \quad (9)$$

Sampling point 0

- There are currently 6 clusters

$$p(z_0 = \text{new} | \vec{z}_{-0}, \vec{x}, \alpha) \propto \frac{0.25}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} 10.00 & 0.00 \\ 10.00 & 0.00 \end{array}, \mathbf{1} \right) = 0.04 \times 0.000000 \quad (3)$$

$$p(z_0 = 1 | \vec{z}_{-0}, \vec{x}, \alpha) \propto \frac{1.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} 10.00 & 4.00 \\ 10.00 & 4.50 \end{array}, \mathbf{1} \right) = 0.16 \times 0.00029 \quad (4)$$

$$p(z_0 = 2 | \vec{z}_{-0}, \vec{x}, \alpha) \propto \frac{1.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} 10.00 & 3.50 \\ 10.00 & 3.00 \end{array}, \mathbf{1} \right) = 0.16 \times 0.00007 \quad (5)$$

$$p(z_0 = 3 | \vec{z}_{-0}, \vec{x}, \alpha) \propto \frac{1.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} 10.00 & -4.50 \\ 10.00 & -5.00 \end{array}, \mathbf{1} \right) = 0.16 \times 0.000000 \quad (6)$$

$$p(z_0 = 4 | \vec{z}_{-0}, \vec{x}, \alpha) \propto \frac{1.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} 10.00 & -2.50 \\ 10.00 & -5.00 \end{array}, \mathbf{1} \right) = 0.16 \times 0.000000 \quad (7)$$

$$p(z_0 = 5 | \vec{z}_{-0}, \vec{x}, \alpha) \propto \frac{1.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} 10.00 & -3.50 \\ 10.00 & -3.50 \end{array}, \mathbf{1} \right) = 0.16 \times 0.000000 \quad (8)$$

$$p(z_0 = 6 | \vec{z}_{-0}, \vec{x}, \alpha) \propto \frac{1.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} 10.00 & 0.50 \\ 10.00 & 0.50 \end{array}, \mathbf{1} \right) = 0.16 \times 0.000000 \quad (9)$$

Sampling point 0

- There are currently 6 clusters

$$p(z_0 = \text{new} | \vec{z}_{-0}, \vec{x}, \alpha) \propto \frac{0.25}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 10.00 & 0.00 \\ 10.00 & 0.00 \end{array} , \mathbf{1} \right) = 0.04 \times 0.000000 \quad (3)$$

$$p(z_0 = 1 | \vec{z}_{-0}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 10.00 & 4.00 \\ 10.00 & 4.50 \end{array} , \mathbf{1} \right) = 0.16 \times 0.00029 \quad (4)$$

$$p(z_0 = 2 | \vec{z}_{-0}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 10.00 & 3.50 \\ 10.00 & 3.00 \end{array} , \mathbf{1} \right) = 0.16 \times 0.00007 \quad (5)$$

$$p(z_0 = 3 | \vec{z}_{-0}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 10.00 & -4.50 \\ 10.00 & -5.00 \end{array} , \mathbf{1} \right) = 0.16 \times 0.000000 \quad (6)$$

$$p(z_0 = 4 | \vec{z}_{-0}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 10.00 & -2.50 \\ 10.00 & -5.00 \end{array} , \mathbf{1} \right) = 0.16 \times 0.000000 \quad (7)$$

$$p(z_0 = 5 | \vec{z}_{-0}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 10.00 & -3.50 \\ 10.00 & -3.50 \end{array} , \mathbf{1} \right) = 0.16 \times 0.000000 \quad (8)$$

$$p(z_0 = 6 | \vec{z}_{-0}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 10.00 & 0.50 \\ 10.00 & 0.50 \end{array} , \mathbf{1} \right) = 0.16 \times 0.000000 \quad (9)$$

- After normalization: {new: 0.00 1: 0.80 2: 0.19 3: 0.00 4: 0.00 5: 0.00 6: 0.00}

Sampling point 0

- There are currently 6 clusters

$$p(z_0 = \text{new} | z_{-0}^{\rightarrow}, \vec{x}, \alpha) \propto \frac{0.25}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 10.00 & 0.00 \\ 10.00 & 0.00 \end{array} , \mathbf{1} \right) = 0.04 \times 0.000000 \quad (3)$$

$$p(z_0 = 1 | z_{-0}^{\rightarrow}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 10.00 & 4.00 \\ 10.00 & 4.50 \end{array} , \mathbf{1} \right) = 0.16 \times 0.00029 \quad (4)$$

$$p(z_0 = 2 | z_{-0}^{\rightarrow}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 10.00 & 3.50 \\ 10.00 & 3.00 \end{array} , \mathbf{1} \right) = 0.16 \times 0.00007 \quad (5)$$

$$p(z_0 = 3 | z_{-0}^{\rightarrow}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 10.00 & -4.50 \\ 10.00 & -5.00 \end{array} , \mathbf{1} \right) = 0.16 \times 0.000000 \quad (6)$$

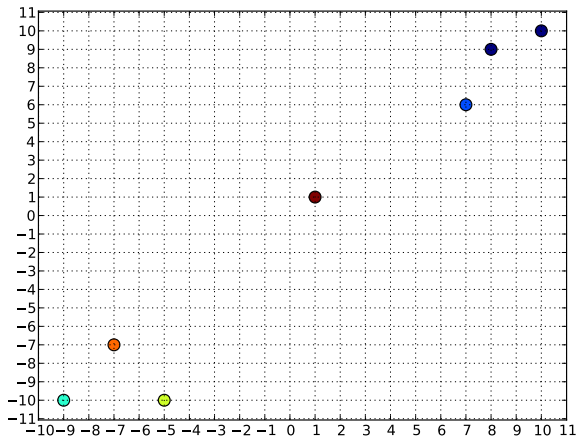
$$p(z_0 = 4 | z_{-0}^{\rightarrow}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 10.00 & -2.50 \\ 10.00 & -5.00 \end{array} , \mathbf{1} \right) = 0.16 \times 0.000000 \quad (7)$$

$$p(z_0 = 5 | z_{-0}^{\rightarrow}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 10.00 & -3.50 \\ 10.00 & -3.50 \end{array} , \mathbf{1} \right) = 0.16 \times 0.000000 \quad (8)$$

$$p(z_0 = 6 | z_{-0}^{\rightarrow}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 10.00 & 0.50 \\ 10.00 & 0.50 \end{array} , \mathbf{1} \right) = 0.16 \times 0.000000 \quad (9)$$

- After normalization: {new: 0.00 1: 0.80 2: 0.19 3: 0.00 4: 0.00 5: 0.00 6: 0.00}
- New assignment = 1

Assignments after sampling point 0



Sampling point 1

Compute the (proportional) probability of assigning data 1 to clusters 1 and 2.

Recall that $\alpha = 0.25$ and

$$p(x | \bar{x}) \propto \exp \left\{ -\sqrt{\left(x_1 - \frac{n}{n+1} \bar{x}_1\right)^2 + \left(x_2 - \frac{n}{n+1} \bar{x}_2\right)^2} \right\} \quad (10)$$

i	x_1	x_2	z_i
0	10	10	1
1	8	9	
2	7	6	2
3	-9	-10	3
4	-5	-10	4
5	-7	-7	5
6	1	1	6

Sampling point 1

- There are currently 6 clusters

$$p(z_1 = 1 | \vec{z}_{-1}, \vec{x}, \alpha) \propto \frac{1.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} 8.00 & 5.00 \\ 9.00 & 5.00 \end{array}, \mathbb{I} \right) = 0.16 \times 0.00674 \quad (10)$$

(11)

Sampling point 1

- There are currently 6 clusters

$$p(z_1 = 1 | \vec{z}_{-1}, \vec{x}, \alpha) \propto \frac{1.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} 8.00 & 5.00 \\ 9.00 & 5.00 \end{array}, \mathbb{I} \right) = 0.16 \times 0.00674 \quad (10)$$

$$p(z_1 = 2 | \vec{z}_{-1}, \vec{x}, \alpha) \propto \frac{1.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} 8.00 & 3.50 \\ 9.00 & 3.00 \end{array}, \mathbb{I} \right) = 0.16 \times 0.00055 \quad (11)$$

(12)

Sampling point 1

- There are currently 6 clusters

$$p(z_1 = \text{new} | \vec{z}_{-1}, \vec{x}, \alpha) \propto \frac{0.25}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 8.00 & 0.00 \\ 9.00 & 0.00 \end{array} , \mathbb{1} \right) = 0.04 \times 0.00001 \quad (10)$$

$$p(z_1 = 1 | \vec{z}_{-1}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 8.00 & 5.00 \\ 9.00 & 5.00 \end{array} , \mathbb{1} \right) = 0.16 \times 0.00674 \quad (11)$$

$$p(z_1 = 2 | \vec{z}_{-1}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 8.00 & 3.50 \\ 9.00 & 3.00 \end{array} , \mathbb{1} \right) = 0.16 \times 0.00055 \quad (12)$$

$$p(z_1 = 3 | \vec{z}_{-1}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 8.00 & -4.50 \\ 9.00 & -5.00 \end{array} , \mathbb{1} \right) = 0.16 \times 0.00000 \quad (13)$$

$$p(z_1 = 4 | \vec{z}_{-1}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 8.00 & -2.50 \\ 9.00 & -5.00 \end{array} , \mathbb{1} \right) = 0.16 \times 0.00000 \quad (14)$$

$$p(z_1 = 5 | \vec{z}_{-1}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 8.00 & -3.50 \\ 9.00 & -3.50 \end{array} , \mathbb{1} \right) = 0.16 \times 0.00000 \quad (15)$$

$$p(z_1 = 6 | \vec{z}_{-1}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 8.00 & 0.50 \\ 9.00 & 0.50 \end{array} , \mathbb{1} \right) = 0.16 \times 0.00001 \quad (16)$$

Sampling point 1

- There are currently 6 clusters

$$p(z_1 = \text{new} | \vec{z}_{-1}, \vec{x}, \alpha) \propto \frac{0.25}{6+0.25} \mathcal{N} \left(\begin{matrix} 8.00 \\ 9.00 \end{matrix} \mid \begin{matrix} 0.00 \\ 0.00 \end{matrix}, \mathbb{1} \right) = 0.04 \times 0.00001 \quad (10)$$

$$p(z_1 = 1 | \vec{z}_{-1}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{matrix} 8.00 \\ 9.00 \end{matrix} \mid \begin{matrix} 5.00 \\ 5.00 \end{matrix}, \mathbb{1} \right) = 0.16 \times 0.00674 \quad (11)$$

$$p(z_1 = 2 | \vec{z}_{-1}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{matrix} 8.00 \\ 9.00 \end{matrix} \mid \begin{matrix} 3.50 \\ 3.00 \end{matrix}, \mathbb{1} \right) = 0.16 \times 0.00055 \quad (12)$$

$$p(z_1 = 3 | \vec{z}_{-1}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{matrix} 8.00 \\ 9.00 \end{matrix} \mid \begin{matrix} -4.50 \\ -5.00 \end{matrix}, \mathbb{1} \right) = 0.16 \times 0.00000 \quad (13)$$

$$p(z_1 = 4 | \vec{z}_{-1}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{matrix} 8.00 \\ 9.00 \end{matrix} \mid \begin{matrix} -2.50 \\ -5.00 \end{matrix}, \mathbb{1} \right) = 0.16 \times 0.00000 \quad (14)$$

$$p(z_1 = 5 | \vec{z}_{-1}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{matrix} 8.00 \\ 9.00 \end{matrix} \mid \begin{matrix} -3.50 \\ -3.50 \end{matrix}, \mathbb{1} \right) = 0.16 \times 0.00000 \quad (15)$$

$$p(z_1 = 6 | \vec{z}_{-1}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{matrix} 8.00 \\ 9.00 \end{matrix} \mid \begin{matrix} 0.50 \\ 0.50 \end{matrix}, \mathbb{1} \right) = 0.16 \times 0.00001 \quad (16)$$

- After normalization: {new: 0.00 1: 0.92 2: 0.08 3: 0.00 4: 0.00 5: 0.00 6: 0.00}

Sampling point 1

- There are currently 6 clusters

$$p(z_1 = \text{new} | \vec{z}_{-1}, \vec{x}, \alpha) \propto \frac{0.25}{6+0.25} \mathcal{N} \left(\begin{matrix} 8.00 \\ 9.00 \end{matrix} \mid \begin{matrix} 0.00 \\ 0.00 \end{matrix}, \mathbb{1} \right) = 0.04 \times 0.00001 \quad (10)$$

$$p(z_1 = 1 | \vec{z}_{-1}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{matrix} 8.00 \\ 9.00 \end{matrix} \mid \begin{matrix} 5.00 \\ 5.00 \end{matrix}, \mathbb{1} \right) = 0.16 \times 0.00674 \quad (11)$$

$$p(z_1 = 2 | \vec{z}_{-1}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{matrix} 8.00 \\ 9.00 \end{matrix} \mid \begin{matrix} 3.50 \\ 3.00 \end{matrix}, \mathbb{1} \right) = 0.16 \times 0.00055 \quad (12)$$

$$p(z_1 = 3 | \vec{z}_{-1}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{matrix} 8.00 \\ 9.00 \end{matrix} \mid \begin{matrix} -4.50 \\ -5.00 \end{matrix}, \mathbb{1} \right) = 0.16 \times 0.00000 \quad (13)$$

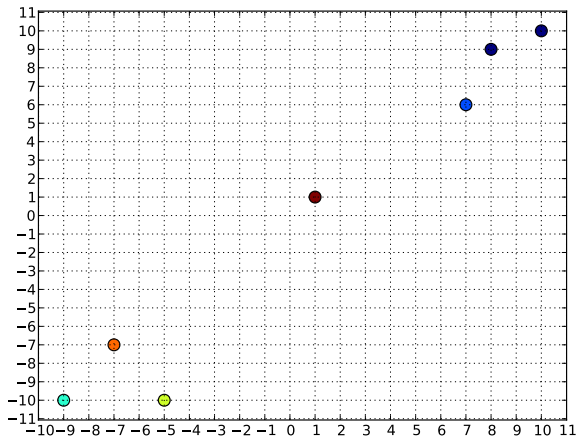
$$p(z_1 = 4 | \vec{z}_{-1}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{matrix} 8.00 \\ 9.00 \end{matrix} \mid \begin{matrix} -2.50 \\ -5.00 \end{matrix}, \mathbb{1} \right) = 0.16 \times 0.00000 \quad (14)$$

$$p(z_1 = 5 | \vec{z}_{-1}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{matrix} 8.00 \\ 9.00 \end{matrix} \mid \begin{matrix} -3.50 \\ -3.50 \end{matrix}, \mathbb{1} \right) = 0.16 \times 0.00000 \quad (15)$$

$$p(z_1 = 6 | \vec{z}_{-1}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{matrix} 8.00 \\ 9.00 \end{matrix} \mid \begin{matrix} 0.50 \\ 0.50 \end{matrix}, \mathbb{1} \right) = 0.16 \times 0.00001 \quad (16)$$

- After normalization: {new: 0.00 1: 0.92 2: 0.08 3: 0.00 4: 0.00 5: 0.00 6: 0.00}
- New assignment = 1

Assignments after sampling point 1



Sampling point 2

Compute the (proportional) probability of assigning data 2 to cluster 1 (but nothing else; there won't be other options).

Recall that $\alpha = 0.25$ and

$$p(x | \bar{x}) \propto \exp \left\{ -\sqrt{\left(x_1 - \frac{n}{n+1} \bar{x}_1\right)^2 + \left(x_2 - \frac{n}{n+1} \bar{x}_2\right)^2} \right\} \quad (17)$$

i	x_1	x_2	z_i
0	10	10	1
1	8	9	1
2	7	6	
3	-9	-10	3
4	-5	-10	4
5	-7	-7	5
6	1	1	6

Sampling point 2

- There are currently 5 clusters

$$p(z_2 = 1 | \vec{z}_{-2}, \vec{x}, \alpha) \propto \frac{2.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} 7.00 & 6.00 \\ 6.00 & 6.33 \end{array}, \mathbf{I} \right) = 0.32 \times 0.34851 \quad (17)$$

(18)

Sampling point 2

- There are currently 5 clusters

$$p(z_2 = \text{new} | \vec{z}_{-2}, \vec{x}, \alpha) \propto \frac{0.25}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 7.00 & 0.00 \\ 6.00 & 0.00 \end{array} \mid \vec{1} \right) = 0.04 \times 0.00010 \quad (17)$$

$$p(z_2 = 1 | \vec{z}_{-2}, \vec{x}, \alpha) \propto \frac{2.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 7.00 & 6.00 \\ 6.00 & 6.33 \end{array} \mid \vec{1} \right) = 0.32 \times 0.34851 \quad (18)$$

$$p(z_2 = 3 | \vec{z}_{-2}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 7.00 & -4.50 \\ 6.00 & -5.00 \end{array} \mid \vec{1} \right) = 0.16 \times 0.00000 \quad (19)$$

$$p(z_2 = 4 | \vec{z}_{-2}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 7.00 & -2.50 \\ 6.00 & -5.00 \end{array} \mid \vec{1} \right) = 0.16 \times 0.00000 \quad (20)$$

$$p(z_2 = 5 | \vec{z}_{-2}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 7.00 & -3.50 \\ 6.00 & -3.50 \end{array} \mid \vec{1} \right) = 0.16 \times 0.00000 \quad (21)$$

$$p(z_2 = 6 | \vec{z}_{-2}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 7.00 & 0.50 \\ 6.00 & 0.50 \end{array} \mid \vec{1} \right) = 0.16 \times 0.00020 \quad (22)$$

Sampling point 2

- There are currently 5 clusters

$$p(z_2 = \text{new} | \vec{z}_{-2}, \vec{x}, \alpha) \propto \frac{0.25}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 7.00 & 0.00 \\ 6.00 & 0.00 \end{array} , \mathbb{1} \right) = 0.04 \times 0.00010 \quad (17)$$

$$p(z_2 = 1 | \vec{z}_{-2}, \vec{x}, \alpha) \propto \frac{2.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 7.00 & 6.00 \\ 6.00 & 6.33 \end{array} , \mathbb{1} \right) = 0.32 \times 0.34851 \quad (18)$$

$$p(z_2 = 3 | \vec{z}_{-2}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 7.00 & -4.50 \\ 6.00 & -5.00 \end{array} , \mathbb{1} \right) = 0.16 \times 0.00000 \quad (19)$$

$$p(z_2 = 4 | \vec{z}_{-2}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 7.00 & -2.50 \\ 6.00 & -5.00 \end{array} , \mathbb{1} \right) = 0.16 \times 0.00000 \quad (20)$$

$$p(z_2 = 5 | \vec{z}_{-2}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 7.00 & -3.50 \\ 6.00 & -3.50 \end{array} , \mathbb{1} \right) = 0.16 \times 0.00000 \quad (21)$$

$$p(z_2 = 6 | \vec{z}_{-2}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 7.00 & 0.50 \\ 6.00 & 0.50 \end{array} , \mathbb{1} \right) = 0.16 \times 0.00020 \quad (22)$$

- After normalization: {new: 0.00 1: 1.00 3: 0.00 4: 0.00 5: 0.00 6: 0.00}

Sampling point 2

- There are currently 5 clusters

$$p(z_2 = \text{new} | \vec{z}_{-2}, \vec{x}, \alpha) \propto \frac{0.25}{6+0.25} \mathcal{N} \left(\begin{matrix} 7.00 \\ 6.00 \end{matrix} \mid \begin{matrix} 0.00 & , \mathbb{1} \\ 0.00 & \end{matrix} \right) = 0.04 \times 0.00010 \quad (17)$$

$$p(z_2 = 1 | \vec{z}_{-2}, \vec{x}, \alpha) \propto \frac{2.00}{6+0.25} \mathcal{N} \left(\begin{matrix} 7.00 \\ 6.00 \end{matrix} \mid \begin{matrix} 6.00 & , \mathbb{1} \\ 6.33 & \end{matrix} \right) = 0.32 \times 0.34851 \quad (18)$$

$$p(z_2 = 3 | \vec{z}_{-2}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{matrix} 7.00 \\ 6.00 \end{matrix} \mid \begin{matrix} -4.50 & , \mathbb{1} \\ -5.00 & \end{matrix} \right) = 0.16 \times 0.00000 \quad (19)$$

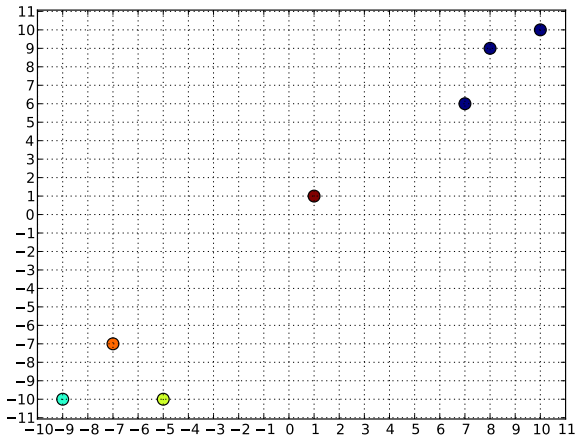
$$p(z_2 = 4 | \vec{z}_{-2}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{matrix} 7.00 \\ 6.00 \end{matrix} \mid \begin{matrix} -2.50 & , \mathbb{1} \\ -5.00 & \end{matrix} \right) = 0.16 \times 0.00000 \quad (20)$$

$$p(z_2 = 5 | \vec{z}_{-2}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{matrix} 7.00 \\ 6.00 \end{matrix} \mid \begin{matrix} -3.50 & , \mathbb{1} \\ -3.50 & \end{matrix} \right) = 0.16 \times 0.00000 \quad (21)$$

$$p(z_2 = 6 | \vec{z}_{-2}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{matrix} 7.00 \\ 6.00 \end{matrix} \mid \begin{matrix} 0.50 & , \mathbb{1} \\ 0.50 & \end{matrix} \right) = 0.16 \times 0.00020 \quad (22)$$

- After normalization: {new: 0.00 1: 1.00 3: 0.00 4: 0.00 5: 0.00 6: 0.00}
- New assignment = 1

Assignments after sampling point 2



Sampling point 3

Compute the (proportional) probability of assigning data 3 to cluster 4 and 5.

Recall that $\alpha = 0.25$ and

$$p(x | \bar{x}) \propto \exp \left\{ -\sqrt{\left(x_1 - \frac{n}{n+1} \bar{x}_1\right)^2 + \left(x_2 - \frac{n}{n+1} \bar{x}_2\right)^2} \right\} \quad (23)$$

i	x_1	x_2	z_i
0	10	10	1
1	8	9	1
2	7	6	1
3	-9	-10	
4	-5	-10	4
5	-7	-7	5
6	1	1	6

Sampling point 3

- There are currently 4 clusters

(23)

Sampling point 3

- There are currently 4 clusters

$$p(z_3 = 4 | \vec{z}_{-3}, \vec{x}, \alpha) \propto \frac{1.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} -9.00 & -2.50 \\ -10.00 & -5.00 \end{array}, \mathbf{1} \right) = 0.16 \times 0.00027 \quad (23)$$

(24)

Sampling point 3

- There are currently 4 clusters

$$p(z_3 = 4 | \vec{z}_{-3}, \vec{x}, \alpha) \propto \frac{1.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} -9.00 & -2.50 \\ -10.00 & -5.00 \end{array}, \mathbb{1} \right) = 0.16 \times 0.00027 \quad (23)$$

$$p(z_3 = 5 | \vec{z}_{-3}, \vec{x}, \alpha) \propto \frac{1.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} -9.00 & -3.50 \\ -10.00 & -3.50 \end{array}, \mathbb{1} \right) = 0.16 \times 0.00020 \quad (24)$$

(25)

Sampling point 3

- There are currently 4 clusters

$$p(z_3 = \text{new} | \vec{z}_{-3}, \vec{x}, \alpha) \propto \frac{0.25}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -9.00 & 0.00 \\ -10.00 & 0.00 \end{array}, \mathbb{1} \right) = 0.04 \times 0.00000 \quad (23)$$

$$p(z_3 = 1 | \vec{z}_{-3}, \vec{x}, \alpha) \propto \frac{3.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -9.00 & 6.25 \\ -10.00 & 6.25 \end{array}, \mathbb{1} \right) = 0.48 \times 0.00000 \quad (24)$$

$$p(z_3 = 4 | \vec{z}_{-3}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -9.00 & -2.50 \\ -10.00 & -5.00 \end{array}, \mathbb{1} \right) = 0.16 \times 0.00027 \quad (25)$$

$$p(z_3 = 5 | \vec{z}_{-3}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -9.00 & -3.50 \\ -10.00 & -3.50 \end{array}, \mathbb{1} \right) = 0.16 \times 0.00020 \quad (26)$$

$$p(z_3 = 6 | \vec{z}_{-3}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -9.00 & 0.50 \\ -10.00 & 0.50 \end{array}, \mathbb{1} \right) = 0.16 \times 0.00000 \quad (27)$$

Sampling point 3

- There are currently 4 clusters

$$p(z_3 = \text{new} | \vec{z}_{-3}, \vec{x}, \alpha) \propto \frac{0.25}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -9.00 & 0.00 \\ -10.00 & 0.00 \end{array}, \mathbb{1} \right) = 0.04 \times 0.00000 \quad (23)$$

$$p(z_3 = 1 | \vec{z}_{-3}, \vec{x}, \alpha) \propto \frac{3.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -9.00 & 6.25 \\ -10.00 & 6.25 \end{array}, \mathbb{1} \right) = 0.48 \times 0.00000 \quad (24)$$

$$p(z_3 = 4 | \vec{z}_{-3}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -9.00 & -2.50 \\ -10.00 & -5.00 \end{array}, \mathbb{1} \right) = 0.16 \times 0.00027 \quad (25)$$

$$p(z_3 = 5 | \vec{z}_{-3}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -9.00 & -3.50 \\ -10.00 & -3.50 \end{array}, \mathbb{1} \right) = 0.16 \times 0.00020 \quad (26)$$

$$p(z_3 = 6 | \vec{z}_{-3}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -9.00 & 0.50 \\ -10.00 & 0.50 \end{array}, \mathbb{1} \right) = 0.16 \times 0.00000 \quad (27)$$

- After normalization: {new: 0.00 1: 0.00 4: 0.58 5: 0.42 6: 0.00}

Sampling point 3

- There are currently 4 clusters

$$p(z_3 = \text{new} | \vec{z}_{-3}, \vec{x}, \alpha) \propto \frac{0.25}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -9.00 & 0.00 \\ -10.00 & 0.00 \end{array} , \mathbb{1} \right) = 0.04 \times 0.00000 \quad (23)$$

$$p(z_3 = 1 | \vec{z}_{-3}, \vec{x}, \alpha) \propto \frac{3.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -9.00 & 6.25 \\ -10.00 & 6.25 \end{array} , \mathbb{1} \right) = 0.48 \times 0.00000 \quad (24)$$

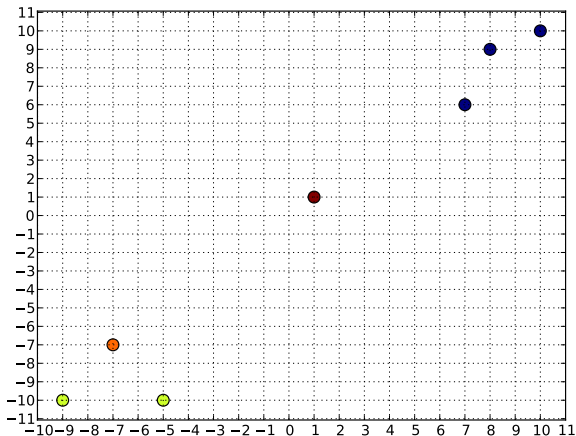
$$p(z_3 = 4 | \vec{z}_{-3}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -9.00 & -2.50 \\ -10.00 & -5.00 \end{array} , \mathbb{1} \right) = 0.16 \times 0.00027 \quad (25)$$

$$p(z_3 = 5 | \vec{z}_{-3}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -9.00 & -3.50 \\ -10.00 & -3.50 \end{array} , \mathbb{1} \right) = 0.16 \times 0.00020 \quad (26)$$

$$p(z_3 = 6 | \vec{z}_{-3}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -9.00 & 0.50 \\ -10.00 & 0.50 \end{array} , \mathbb{1} \right) = 0.16 \times 0.00000 \quad (27)$$

- After normalization: {new: 0.00 1: 0.00 4: 0.58 5: 0.42 6: 0.00}
- New assignment = 4

Assignments after sampling point 3



Sampling point 4

Compute the (proportional) probability of assigning data 4 to cluster 4 and 5.

Recall that $\alpha = 0.25$ and

$$p(x | \bar{x}) \propto \exp \left\{ -\sqrt{\left(x_1 - \frac{n}{n+1} \bar{x}_1\right)^2 + \left(x_2 - \frac{n}{n+1} \bar{x}_2\right)^2} \right\} \quad (28)$$

i	x_1	x_2	z_i
0	10	10	1
1	8	9	1
2	7	6	1
3	-9	-10	4
4	-5	-10	
5	-7	-7	5
6	1	1	6

Sampling point 4

- There are currently 4 clusters

$$p(z_4 = 4 | \vec{z}_{-4}, \vec{x}, \alpha) \propto \frac{1.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} -5.00 & -4.50 \\ -10.00 & -5.00 \end{array}, \mathbf{1} \right) = 0.16 \times 0.00657 \quad (28)$$

(29)

Sampling point 4

- There are currently 4 clusters

$$p(z_4 = 4 | \vec{z}_{-4}, \vec{x}, \alpha) \propto \frac{1.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} -5.00 & -4.50 \\ -10.00 & -5.00 \end{array}, \mathbb{1} \right) = 0.16 \times 0.00657 \quad (28)$$

$$p(z_4 = 5 | \vec{z}_{-4}, \vec{x}, \alpha) \propto \frac{1.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} -5.00 & -3.50 \\ -10.00 & -3.50 \end{array}, \mathbb{1} \right) = 0.16 \times 0.00127 \quad (29)$$

(30)

Sampling point 4

- There are currently 4 clusters

$$p(z_4 = \text{new} | \vec{z}_{-4}, \vec{x}, \alpha) \propto \frac{0.25}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -5.00 & 0.00 \\ -10.00 & 0.00 \end{array} , \mathbb{1} \right) = 0.04 \times 0.00001 \quad (28)$$

$$p(z_4 = 1 | \vec{z}_{-4}, \vec{x}, \alpha) \propto \frac{3.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -5.00 & 6.25 \\ -10.00 & 6.25 \end{array} , \mathbb{1} \right) = 0.48 \times 0.00000 \quad (29)$$

$$p(z_4 = 4 | \vec{z}_{-4}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -5.00 & -4.50 \\ -10.00 & -5.00 \end{array} , \mathbb{1} \right) = 0.16 \times 0.00657 \quad (30)$$

$$p(z_4 = 5 | \vec{z}_{-4}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -5.00 & -3.50 \\ -10.00 & -3.50 \end{array} , \mathbb{1} \right) = 0.16 \times 0.00127 \quad (31)$$

$$p(z_4 = 6 | \vec{z}_{-4}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -5.00 & 0.50 \\ -10.00 & 0.50 \end{array} , \mathbb{1} \right) = 0.16 \times 0.00001 \quad (32)$$

Sampling point 4

- There are currently 4 clusters

$$p(z_4 = \text{new} | \vec{z}_{-4}, \vec{x}, \alpha) \propto \frac{0.25}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -5.00 & 0.00 \\ -10.00 & 0.00 \end{array}, \mathbb{1} \right) = 0.04 \times 0.00001 \quad (28)$$

$$p(z_4 = 1 | \vec{z}_{-4}, \vec{x}, \alpha) \propto \frac{3.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -5.00 & 6.25 \\ -10.00 & 6.25 \end{array}, \mathbb{1} \right) = 0.48 \times 0.00000 \quad (29)$$

$$p(z_4 = 4 | \vec{z}_{-4}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -5.00 & -4.50 \\ -10.00 & -5.00 \end{array}, \mathbb{1} \right) = 0.16 \times 0.00657 \quad (30)$$

$$p(z_4 = 5 | \vec{z}_{-4}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -5.00 & -3.50 \\ -10.00 & -3.50 \end{array}, \mathbb{1} \right) = 0.16 \times 0.00127 \quad (31)$$

$$p(z_4 = 6 | \vec{z}_{-4}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -5.00 & 0.50 \\ -10.00 & 0.50 \end{array}, \mathbb{1} \right) = 0.16 \times 0.00001 \quad (32)$$

- After normalization: {new: 0.00 1: 0.00 4: 0.84 5: 0.16 6: 0.00}

Sampling point 4

- There are currently 4 clusters

$$p(z_4 = \text{new} | \vec{z}_{-4}, \vec{x}, \alpha) \propto \frac{0.25}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -5.00 & 0.00 \\ -10.00 & 0.00 \end{array}, \mathbb{1} \right) = 0.04 \times 0.00001 \quad (28)$$

$$p(z_4 = 1 | \vec{z}_{-4}, \vec{x}, \alpha) \propto \frac{3.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -5.00 & 6.25 \\ -10.00 & 6.25 \end{array}, \mathbb{1} \right) = 0.48 \times 0.00000 \quad (29)$$

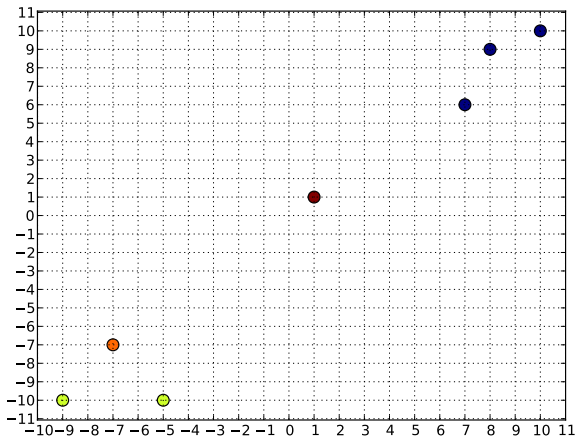
$$p(z_4 = 4 | \vec{z}_{-4}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -5.00 & -4.50 \\ -10.00 & -5.00 \end{array}, \mathbb{1} \right) = 0.16 \times 0.00657 \quad (30)$$

$$p(z_4 = 5 | \vec{z}_{-4}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -5.00 & -3.50 \\ -10.00 & -3.50 \end{array}, \mathbb{1} \right) = 0.16 \times 0.00127 \quad (31)$$

$$p(z_4 = 6 | \vec{z}_{-4}, \vec{x}, \alpha) \propto \frac{1.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} -5.00 & 0.50 \\ -10.00 & 0.50 \end{array}, \mathbb{1} \right) = 0.16 \times 0.00001 \quad (32)$$

- After normalization: {new: 0.00 1: 0.00 4: 0.84 5: 0.16 6: 0.00}
- New assignment = 4

Assignments after sampling point 4



Sampling point 5

Compute the (proportional) probability of assigning data 5 to cluster 4 (but nothing else is viable).

Recall that $\alpha = 0.25$ and

$$p(x | \bar{x}) \propto \exp \left\{ -\sqrt{\left(x_1 - \frac{n}{n+1} \bar{x}_1\right)^2 + \left(x_2 - \frac{n}{n+1} \bar{x}_2\right)^2} \right\} \quad (33)$$

i	x_1	x_2	z_i
0	10	10	1
1	8	9	1
2	7	6	1
3	-9	-10	4
4	-5	-10	4
5	-7	-7	
6	1	1	6

Sampling point 5

- There are currently 3 clusters

$$p(z_5 = 4 | \vec{z}_{-5}, \vec{x}, \alpha) \propto \frac{2.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} -7.00 & -4.67 \\ -7.00 & -6.67 \end{array}, \mathbf{1} \right) = 0.32 \times 0.09470 \quad (33)$$

(34)

Sampling point 5

- There are currently 3 clusters

$$p(z_5 = \text{new} | \vec{z}_{-5}, \vec{x}, \alpha) \propto \frac{0.25}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} -7.00 & 0.00 \\ -7.00 & 0.00 \end{array} , \mathbf{1} \right) = 0.04 \times 0.00005 \quad (33)$$

$$p(z_5 = 1 | \vec{z}_{-5}, \vec{x}, \alpha) \propto \frac{3.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} -7.00 & 6.25 \\ -7.00 & 6.25 \end{array} , \mathbf{1} \right) = 0.48 \times 0.00000 \quad (34)$$

$$p(z_5 = 4 | \vec{z}_{-5}, \vec{x}, \alpha) \propto \frac{2.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} -7.00 & -4.67 \\ -7.00 & -6.67 \end{array} , \mathbf{1} \right) = 0.32 \times 0.09470 \quad (35)$$

$$p(z_5 = 6 | \vec{z}_{-5}, \vec{x}, \alpha) \propto \frac{1.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} -7.00 & 0.50 \\ -7.00 & 0.50 \end{array} , \mathbf{1} \right) = 0.16 \times 0.00002 \quad (36)$$

Sampling point 5

- There are currently 3 clusters

$$p(z_5 = \text{new} | z_{-5}^{\rightarrow}, \vec{x}, \alpha) \propto \frac{0.25}{6 + 0.25} \mathcal{N} \left(\begin{matrix} -7.00 \\ -7.00 \end{matrix} \mid \begin{matrix} 0.00 \\ 0.00 \end{matrix}, \mathbf{1} \right) = 0.04 \times 0.00005 \quad (33)$$

$$p(z_5 = 1 | z_{-5}^{\rightarrow}, \vec{x}, \alpha) \propto \frac{3.00}{6 + 0.25} \mathcal{N} \left(\begin{matrix} -7.00 \\ -7.00 \end{matrix} \mid \begin{matrix} 6.25 \\ 6.25 \end{matrix}, \mathbf{1} \right) = 0.48 \times 0.00000 \quad (34)$$

$$p(z_5 = 4 | z_{-5}^{\rightarrow}, \vec{x}, \alpha) \propto \frac{2.00}{6 + 0.25} \mathcal{N} \left(\begin{matrix} -7.00 \\ -7.00 \end{matrix} \mid \begin{matrix} -4.67 \\ -6.67 \end{matrix}, \mathbf{1} \right) = 0.32 \times 0.09470 \quad (35)$$

$$p(z_5 = 6 | z_{-5}^{\rightarrow}, \vec{x}, \alpha) \propto \frac{1.00}{6 + 0.25} \mathcal{N} \left(\begin{matrix} -7.00 \\ -7.00 \end{matrix} \mid \begin{matrix} 0.50 \\ 0.50 \end{matrix}, \mathbf{1} \right) = 0.16 \times 0.00002 \quad (36)$$

- After normalization: {new: 0.00 1: 0.00 4: 1.00 6: 0.00}

Sampling point 5

- There are currently 3 clusters

$$p(z_5 = \text{new} | \vec{z}_{-5}, \vec{x}, \alpha) \propto \frac{0.25}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} -7.00 & 0.00 \\ -7.00 & 0.00 \end{array} , \mathbf{1} \right) = 0.04 \times 0.00005 \quad (33)$$

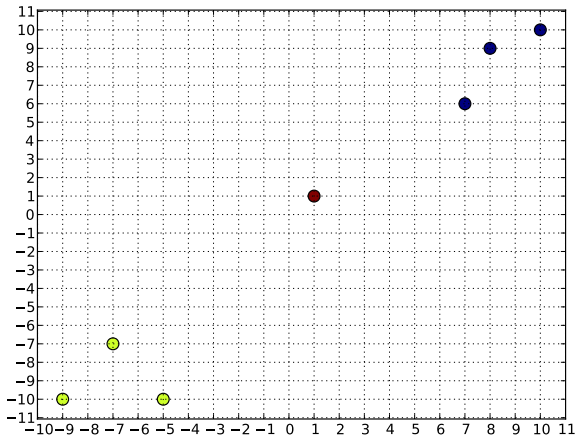
$$p(z_5 = 1 | \vec{z}_{-5}, \vec{x}, \alpha) \propto \frac{3.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} -7.00 & 6.25 \\ -7.00 & 6.25 \end{array} , \mathbf{1} \right) = 0.48 \times 0.00000 \quad (34)$$

$$p(z_5 = 4 | \vec{z}_{-5}, \vec{x}, \alpha) \propto \frac{2.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} -7.00 & -4.67 \\ -7.00 & -6.67 \end{array} , \mathbf{1} \right) = 0.32 \times 0.09470 \quad (35)$$

$$p(z_5 = 6 | \vec{z}_{-5}, \vec{x}, \alpha) \propto \frac{1.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} -7.00 & 0.50 \\ -7.00 & 0.50 \end{array} , \mathbf{1} \right) = 0.16 \times 0.00002 \quad (36)$$

- After normalization: {new: 0.00 1: 0.00 4: 1.00 6: 0.00}
- New assignment = 4

Assignments after sampling point 5



Sampling point 6

Compute the (proportional) probability of assigning data 6 to a new cluster and cluster 1.

Recall that $\alpha = 0.25$ and

$$p(x | \bar{x}) \propto \exp \left\{ -\sqrt{\left(x_1 - \frac{n}{n+1} \bar{x}_1\right)^2 + \left(x_2 - \frac{n}{n+1} \bar{x}_2\right)^2} \right\} \quad (37)$$

i	x_1	x_2	z_i
0	10	10	1
1	8	9	1
2	7	6	1
3	-9	-10	4
4	-5	-10	4
5	-7	-7	4
6	1	1	

Sampling point 6

- There are currently 2 clusters

$$p(z_6 = \text{new} | \vec{z}_{-6}, \vec{x}, \alpha) \propto \frac{0.25}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} 1.00 & 0.00 \\ 1.00 & 0.00 \end{array}, \mathbf{1} \right) = 0.04 \times 0.24312 \quad (37)$$

(38)

Sampling point 6

- There are currently 2 clusters

$$p(z_6 = \text{new} | \vec{z}_{-6}, \vec{x}, \alpha) \propto \frac{0.25}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} 1.00 & 0.00 \\ 1.00 & 0.00 \end{array}, \mathbb{I} \right) = 0.04 \times 0.24312 \quad (37)$$

$$p(z_6 = 1 | \vec{z}_{-6}, \vec{x}, \alpha) \propto \frac{3.00}{6 + 0.25} \mathcal{N} \left(\begin{array}{c|c} 1.00 & 6.25 \\ 1.00 & 6.25 \end{array}, \mathbb{I} \right) = 0.48 \times 0.00060 \quad (38)$$

(39)

Sampling point 6

- There are currently 2 clusters

$$p(z_6 = \text{new} | \vec{z}_{-6}, \vec{x}, \alpha) \propto \frac{0.25}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 1.00 & 0.00 \\ 1.00 & 0.00 \end{array}, \mathbb{1} \right) = 0.04 \times 0.24312 \quad (37)$$

$$p(z_6 = 1 | \vec{z}_{-6}, \vec{x}, \alpha) \propto \frac{3.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 1.00 & 6.25 \\ 1.00 & 6.25 \end{array}, \mathbb{1} \right) = 0.48 \times 0.00060 \quad (38)$$

$$p(z_6 = 4 | \vec{z}_{-6}, \vec{x}, \alpha) \propto \frac{3.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 1.00 & -5.25 \\ 1.00 & -6.75 \end{array}, \mathbb{1} \right) = 0.48 \times 0.00005 \quad (39)$$

Sampling point 6

- There are currently 2 clusters

$$p(z_6 = \text{new} | \vec{z}_{-6}, \vec{x}, \alpha) \propto \frac{0.25}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 1.00 & 0.00 \\ 1.00 & 0.00 \end{array}, \mathbb{1} \right) = 0.04 \times 0.24312 \quad (37)$$

$$p(z_6 = 1 | \vec{z}_{-6}, \vec{x}, \alpha) \propto \frac{3.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 1.00 & 6.25 \\ 1.00 & 6.25 \end{array}, \mathbb{1} \right) = 0.48 \times 0.00060 \quad (38)$$

$$p(z_6 = 4 | \vec{z}_{-6}, \vec{x}, \alpha) \propto \frac{3.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 1.00 & -5.25 \\ 1.00 & -6.75 \end{array}, \mathbb{1} \right) = 0.48 \times 0.00005 \quad (39)$$

- After normalization: {new: 0.97 1: 0.03 4: 0.00}

Sampling point 6

- There are currently 2 clusters

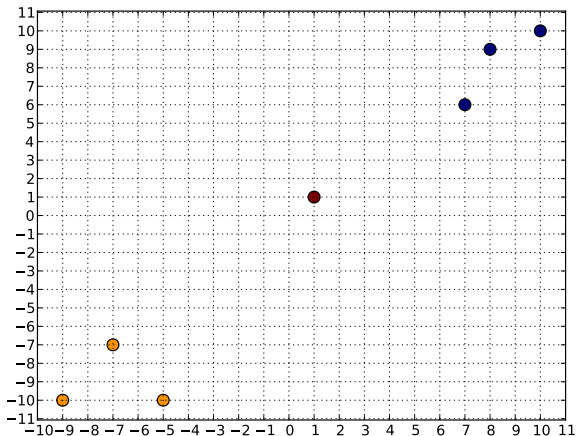
$$p(z_6 = \text{new} | \vec{z}_{-6}, \vec{x}, \alpha) \propto \frac{0.25}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 1.00 & 0.00 \\ 1.00 & 0.00 \end{array}, \mathbb{1} \right) = 0.04 \times 0.24312 \quad (37)$$

$$p(z_6 = 1 | \vec{z}_{-6}, \vec{x}, \alpha) \propto \frac{3.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 1.00 & 6.25 \\ 1.00 & 6.25 \end{array}, \mathbb{1} \right) = 0.48 \times 0.00060 \quad (38)$$

$$p(z_6 = 4 | \vec{z}_{-6}, \vec{x}, \alpha) \propto \frac{3.00}{6+0.25} \mathcal{N} \left(\begin{array}{c|c} 1.00 & -5.25 \\ 1.00 & -6.75 \end{array}, \mathbb{1} \right) = 0.48 \times 0.00005 \quad (39)$$

- After normalization: {new: 0.97 1: 0.03 4: 0.00}
- New assignment = 0

Assignments after sampling point 6



But this is a course about text!

- Base distribution can be any distribution
- Including multinomial!
- Generalizes nicely to documents as well
- (But intuition easier with continuous case)