

An Example of 2 Clusters

Random Initialization

$$p(\theta_1) = p(\theta_2) = 0.5$$

	$p(w \theta_1)$	$p(w \theta_2)$
text	0.5	0.1
mining	0.2	0.1
medical	0.2	0.75
health	0.1	0.05

E-step Document d

Hidden variables:

$$Z_d \in \{1, 2\}$$

	$c(w,d)$
text	2
mining	2
medical	0
health	0

$$\begin{aligned} p(Z_d = 1 | d) &= \frac{p(\theta_1) p(\text{"text"} | \theta_1)^2 p(\text{"mining"} | \theta_1)^2}{p(\theta_1) p(\text{"text"} | \theta_1)^2 p(\text{"mining"} | \theta_1)^2 + p(\theta_2) p(\text{"text"} | \theta_2)^2 p(\text{"mining"} | \theta_2)^2} \\ &= \frac{0.5 * 0.5^2 * 0.2^2}{0.5 * 0.5^2 * 0.2^2 + 0.5 * 0.1^2 * 0.1^2} = \frac{100}{101} \end{aligned}$$

$$p(Z_d = 2 | d) = ?$$