The Expectation-Maximization (EM) Algorithm

Hidden Variable:
$z \in \{0, 1\}$

	z
the	1
paper	1
presents	1
a	1
text	0
mining	0
algorithm	0
for	1
clustering	0
...	...

Initialize $p(w|\theta_d)$ with random values.
Then iteratively improve it using E-step & M-step.
Stop when likelihood doesn't change.

$$p^{(n)}(z=0 \mid w) = \frac{p(\theta_d)p^{(n)}(w \mid \theta_d)}{p(\theta_d)p^{(n)}(w \mid \theta_d) + p(\theta_B)p(w \mid \theta_B)}$$

E-step

How likely w is from θ_d

$$p^{(n+1)}(w \mid \theta_d) = \frac{c(w,d)p^{(n)}(z=0 \mid w)}{\sum_{w' \in V} c(w',d)p^{(n)}(z=0 \mid w')}$$

M-step