베이즈 가우시안 혼합 모델(Bayes Gaussian Mixture model)

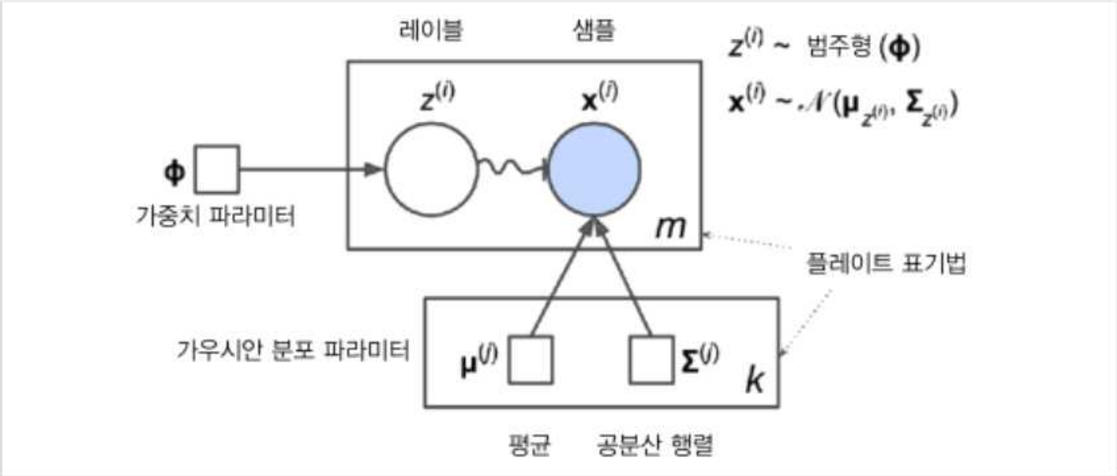
- 사전 지식(*ex 클러스터가 적을 것이라는 믿음 등*)을 반영한 GMM모델
- 최적의 클러스터 개수를 수동으로 찾지 않고 불필요한 가중치를 0으로(또는 0에 가깝게) 만듦
- 데이터의 수가 많아질수록 사전 지식은 중요하지 않음

<기존 GMM 과의 차이>

- 기존 GMM

\- Σ, μ, π -> 파라미터

\- z : 클러스터 할당 변수



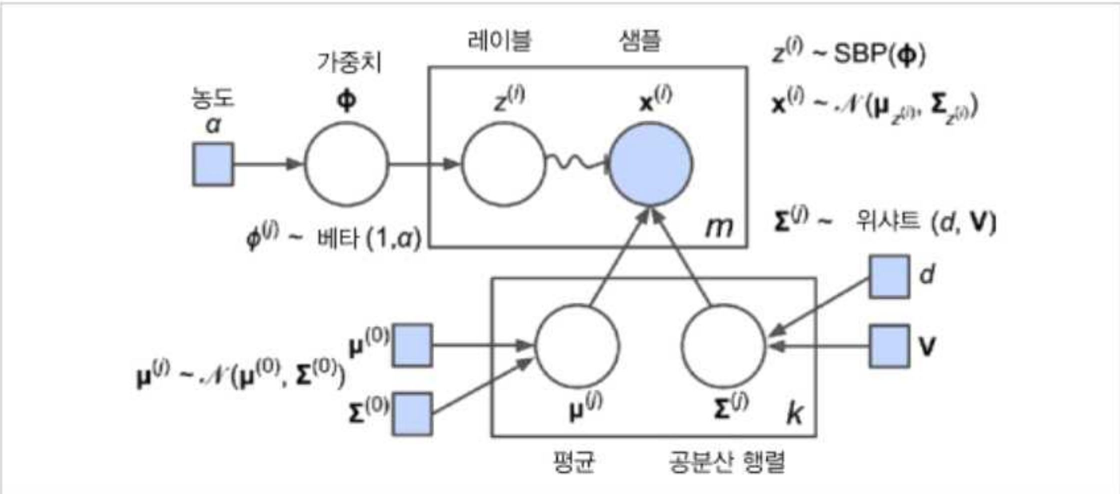
<GMM, 출처 : 핸즈온 머신러닝2>

- Bayesian Gaussian Mixture model

\- Σ, μ, π -> 잠재변수

\- z : 클러스터 파라미터, 할당 변수, 파라미터에 의해 사전 믿음이 반영됨.

\- λ : 파라미터



<Bayesian Gaussian Mixture model, 출처 : 핸즈온 머신러닝2>

<Train>

- w에 대한 prior(*π는 베타분포에서 샘플링, σ는 공분산 행렬에서 샘플링)가 있기 때문에 posterior을 구하겠다.

$posterior = p(z|X) = \frac{p(X|z)p(z)}{evidence}, evidence = \int p(X|z)p(z)dz$-> evidence를 구하기 어려움

$=> q(z|\lambda)$

- 목적 : $p(z|X)$와 $q(z|\lambda)$를 가장 비슷하게 하는 λ를 찾는 것이 목표
- $D_{KL}(q(z|\lambda)\|p(z|X)) = \sum_z q(z|\lambda)\ln(\frac{q(z|\lambda)}{p(z|X)})$

$$
\begin{aligned}
&= E_q(\ln(\frac{q(z|\lambda)}{p(z|X)})) \\
&= E_q(\ln(q(z|\lambda)) - E_q(\ln p(z|X)) \\
&= E_q(\ln(q(z|\lambda)) - E_q(\ln p(z,X)) + E_q(\ln p(X)) \\
&= \ln p(X) - (E_q(\ln p(z,X)) - E_q(\ln q(z|\lambda))) \\
&= \ln p(X|\lambda) - ELBO, ELBO = E_q(\ln p(z,X)) - E_q(\ln q(z|\lambda))
\end{aligned}
$$

**ELBO는 evidence의 lower bound*

=> $\therefore$ **ELBO를 최대화하는 λ를 찾음**(EM Algorithm / Gradient descent 사용)