Week 1.

Gaussian Model Learning

1.4.3. EM Algorithm [Advanced]

Robotics

Estimation and Learning with Dan Lee

Week 1. Gaussian Model Learning

1.4.3 EM Algorithm [Advanced]

Penn Engineering

Expectation-Maximization (EM)

- EM as lower-bound maximization

$$\arg\max_{\theta} \sum_{i} \ln p(x_i|\theta)$$

θ : All parameters

(1) Jensen's inequality

(2) Latent variable and marginal probability

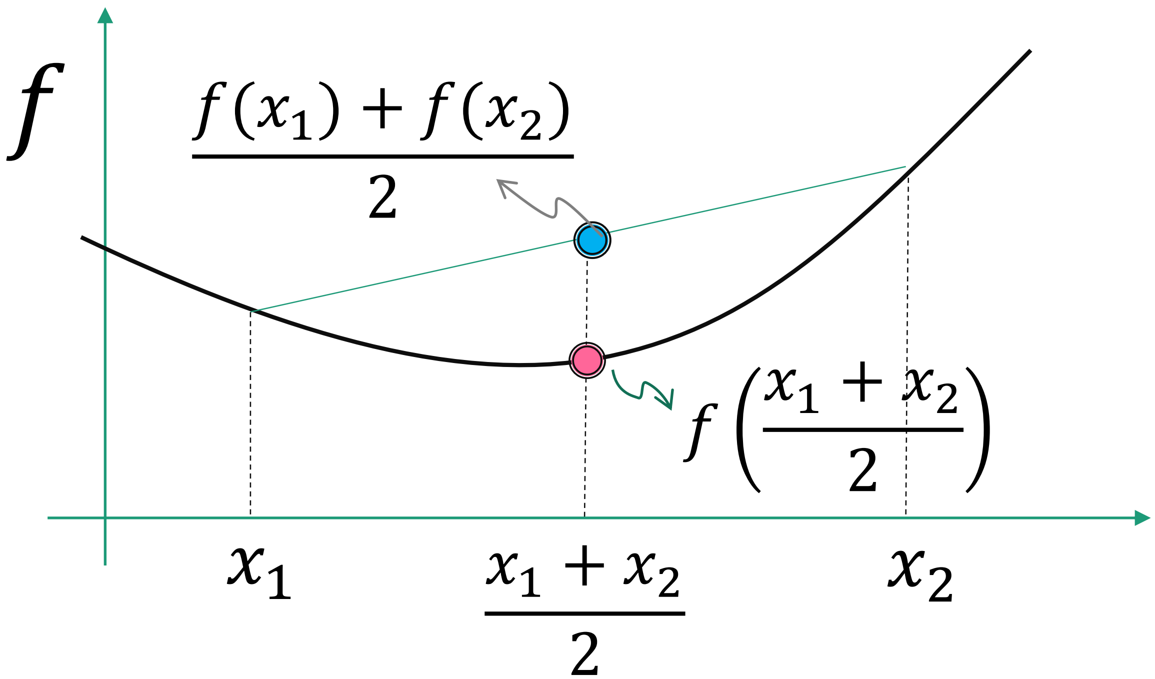
(3) Procedure : E-step and M-step.

Expectation-Maximization (EM)

(1) Jensen's inequality

f : **convex** function

$$\boxed{f\left(\frac{x_1+x_2}{2}\right) \leq \frac{f(x_1)+f(x_2)}{2}}$$



Expectation-Maximization (EM)

(1) Jensen's inequality

f : **convex** function

$$f\left(\sum a_i x_i\right) \leq \sum a_i f(x_i)$$

$$\sum a_i = 1$$

$$a_i \geq 0$$

Expectation-Maximization (EM)

(1) Jensen's inequality

f : **concave** function

$$f\left(\sum a_i x_i\right) \geq \sum a_i f(x_i)$$

$$\sum a_i = 1$$

$$a_i \geq 0$$

Expectation-Maximization (EM)

(1) Jensen's inequality

$\ln$ is *concave*

$$\ln\left(\sum a_i p_i\right) \geq \sum a_i \ln p_i$$

$$\sum a_i = 1$$

$$a_i \geq 0$$

Expectation-Maximization (EM)

(2) latent variable z

$$p(X|\theta) = \sum_{Z} p(X, Z|\theta)$$

(From definition of marginal probability)

Expectation-Maximization (EM)

(2) latent variable

$$\ln p(X|\theta) = \ln \sum_{Z} p(X, Z|\theta)$$

Log-likelihood

$$= \ln \sum_{Z} q(Z) \frac{p(X, Z|\theta)}{q(Z)} \geq \sum_{Z} q(Z) \ln \frac{p(X, Z|\theta)}{q(Z)}$$

Lower bound

Note: $q(Z)$ is a valid probability distribution over Z.

Expectation-Maximization (EM)

(2) latent variable

$$\ln p(X|\theta) = \ln \sum_{Z} p(X,Z|\theta) \geq \sum_{Z} q(Z) \ln \frac{p(X,Z|\theta)}{q(Z)}$$

Log-likelihood

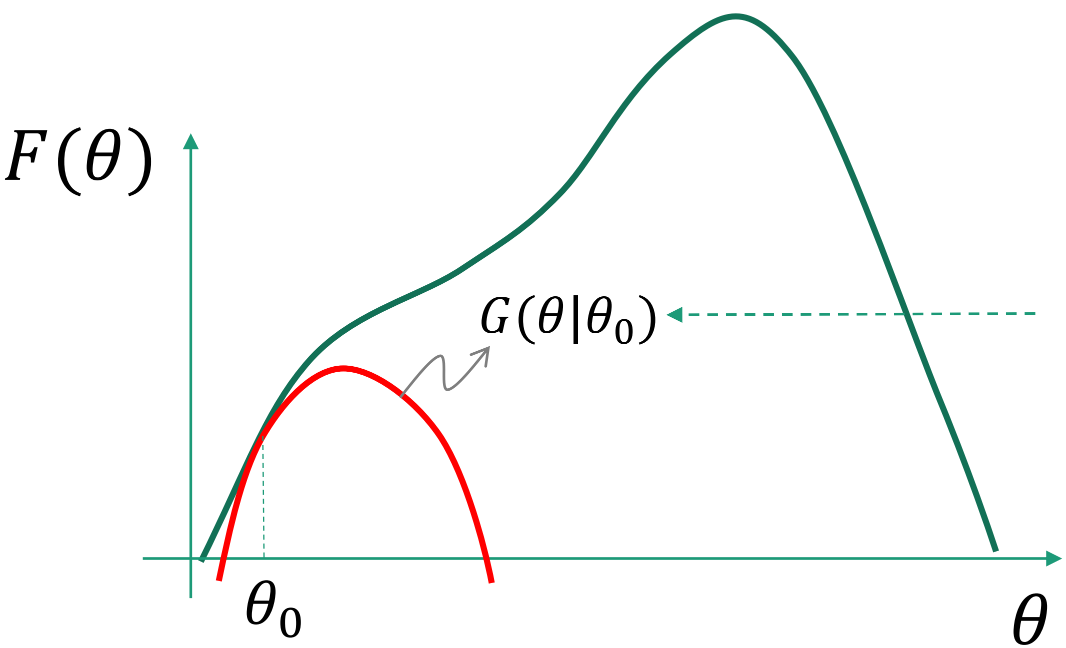
Lower bound

Find q!

Expectation-Maximization (EM)

(3a) Find a lower bound G with an initial guess

$$\underbrace{\ln p(X|\theta)}_{F} \geq \underbrace{\sum_{Z} q(Z)\ln\frac{p(X,Z|\theta)}{q(Z)}}_{G}$$

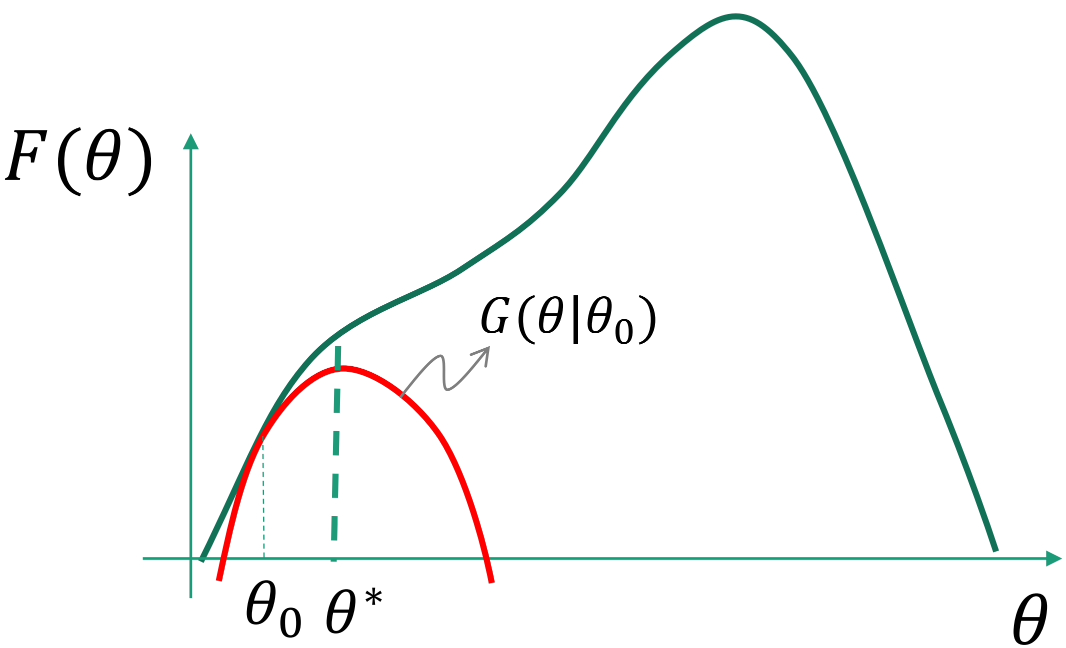


This Implies that $q(Z)$ is found based on θ_0.

Expectation-Maximization (EM)

(3b) Find $\theta^* = \arg\max_{\theta} G(\theta|\theta_0)$

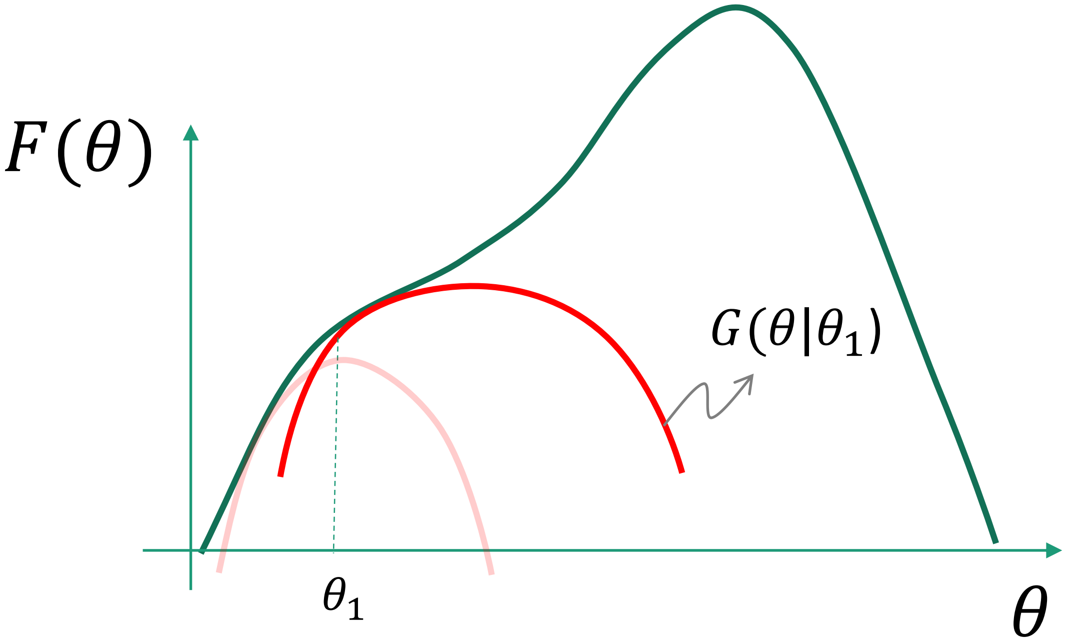
$$\underbrace{\ln p(X|\theta)}_{F} \geq \underbrace{\sum_Z q(Z) \ln \frac{p(X,Z|\theta)}{q(Z)}}_{G}$$



Expectation-Maximization (EM)

(3c) Find a new lower bound G with $\theta_1 \leftarrow \theta^*$

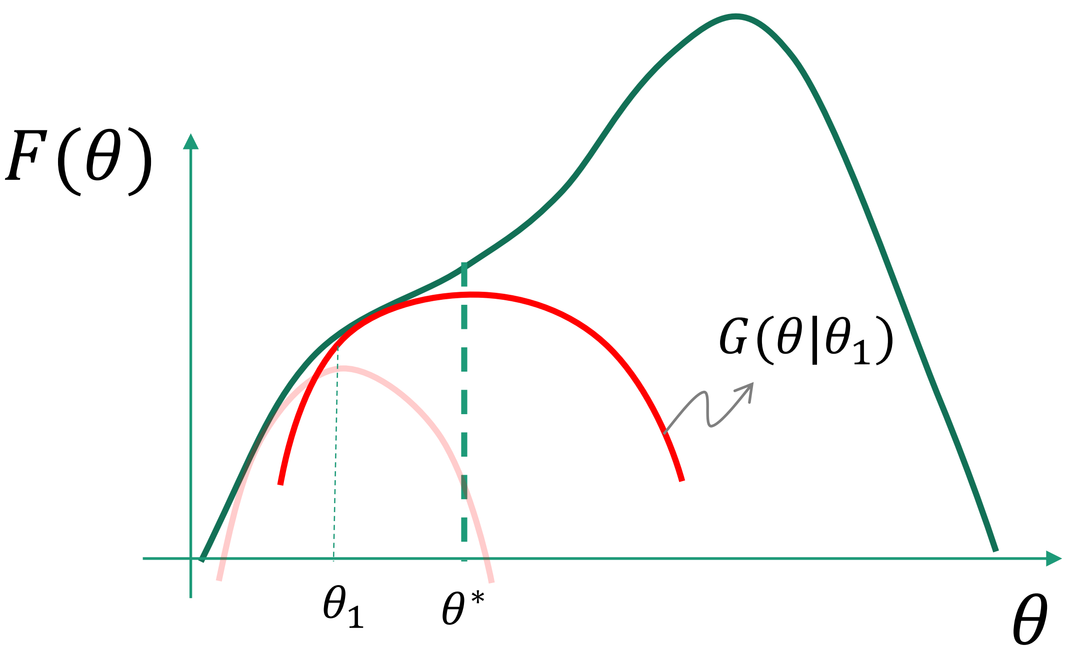
$$\underbrace{\ln p(X|\theta)}_{F} \geq \underbrace{\sum_Z q(Z) \ln \frac{p(X,Z|\theta)}{q(Z)}}_{G}$$



Expectation-Maximization (EM)

(3d) Find $\theta^* = \arg\max_{\theta} G(\theta|\theta_1)$

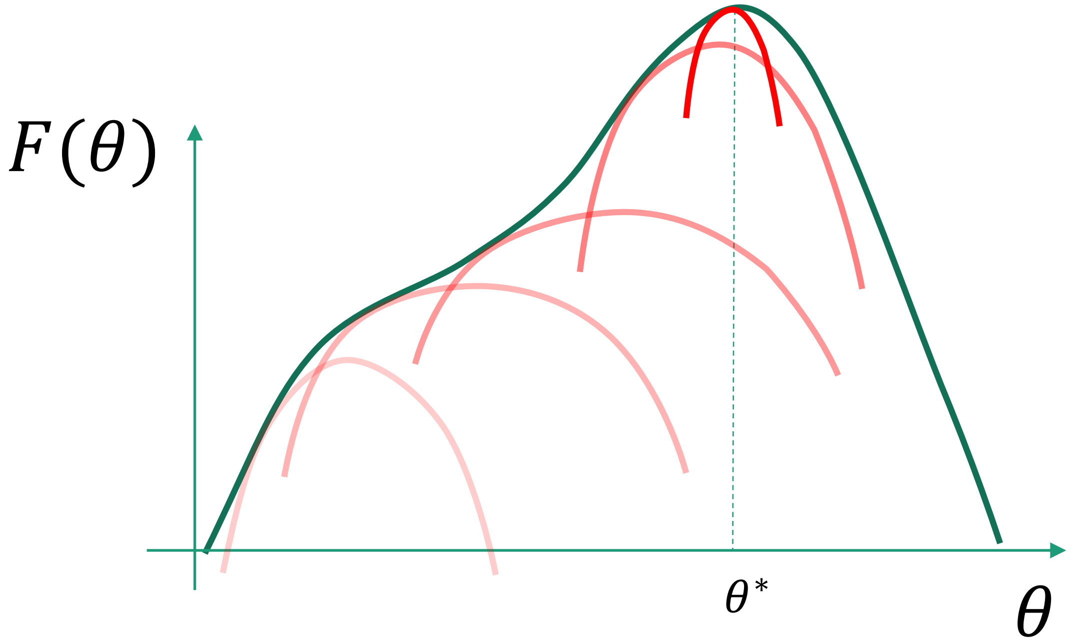
$$\underbrace{\ln p(X|\theta)}_{F} \geq \underbrace{\sum_{Z} q(Z) \ln \frac{p(X,Z|\theta)}{q(Z)}}_{G}$$



Expectation-Maximization (EM)

(4e) Repeat (until converged)

$$\underbrace{\ln p(X|\theta)}_{F} \geq \underbrace{\sum_{Z} q(Z) \ln \frac{p(X,Z|\theta)}{q(Z)}}_{G}$$



Expectation-Maximization (EM)

Initialize $\hat{\theta}$

E-step

Find a lower bound $G\left(\theta \middle| \hat{\theta}\right)$

This Implies that $q(Z)$ is found based on $\hat{\theta}$.

$$\hat{\theta}_{new} = \arg\max_{\theta} G(\theta|\hat{\theta})$$

Stop if converged

Expectation-Maximization (EM)