Expectation-Maximization – EM Algorithm Solved Example

- Expectation-Maximization (EM) – a very popular technique for estimating parameters of probabilistic models.
- Many popular algorithms like Hidden Markov Models, Gaussian Mixtures, Kalman Filters, and others uses EM technique.
- It is beneficial **when working with data that is incomplete, has missing data points, or has unobserved latent variables**.

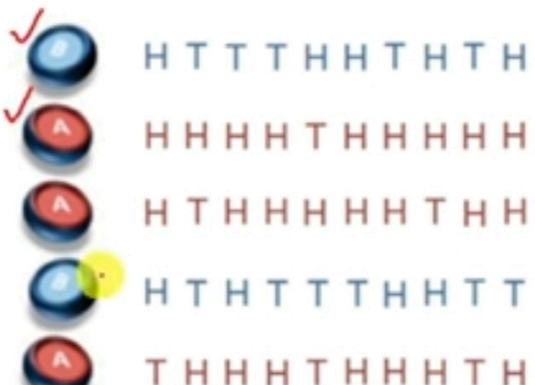
Expectation-Maximization – EM Algorithm Solved Example

- Assume that we have two coins, C1 and C2
- Assume the bias of C_1 is θ_1 (i.e., probability of getting heads with C_1)
- Assume the bias of C_2 is θ_2 (i.e., probability of getting heads with C2)
- We want to find θ_1, θ_2 by performing a number of trials (i.e., coin tosses)

Expectation-Maximization – EM Algorithm Solved Example

First experiment

- We choose 5 times one of the coins.
- We toss the chosen coin 10 times



Expectation-Maximization – EM Algorithm Solved Example

First experiment

- We choose 5 times one of the coins.
- We toss the chosen coin 10 times

H T T T H H T H T H

H H H H T H H H H H

H T H H H H H T H H

H T H T T T H H T T

T H H H T H H H T H

$$\theta_1 = \frac{number\ of\ heads\ using\ C1}{total\ number\ of\ flips\ using\ C1}$$

$$\theta_2 = \frac{number\ of\ heads\ using\ C2}{total\ number\ of\ flips\ using\ C2}$$

Expectation-Maximization – EM Algorithm Solved Example

H T T T H H T H T H

H H H H T H H H H H

H T H H H H H T H H

H T H T T T H H T T

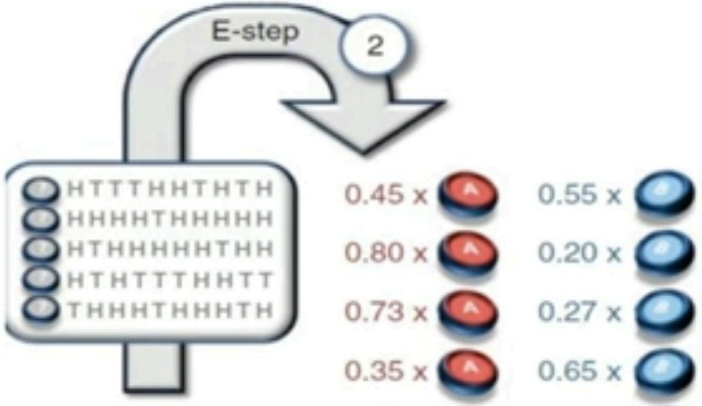
T H H H T H H H T H

Coin A	Coin B
	5 H, 5 T
9 H, 1 T	
8 H, 2 T	
	4 H, 6 T
7 H, 3 T	
24 H, 6 T	9 H, 11 T

$$\theta_1 = \frac{24}{24+6} = 0.8$$

$$\theta_2 = \frac{9}{9+11} = 0.45$$

Expectation-Maximization – EM Algorithm Solved Example



Coin A	Coin B
≈ 2.2 H, 2.2 T	≈ 2.8 H, 2.8 T
≈ 7.2 H, 0.8 T	≈ 1.8 H, 0.2 T
≈ 5.9 H, 1.5 T	≈ 2.1 H, 0.5 T
≈ 1.4 H, 2.1 T	≈ 2.6 H, 3.9 T
≈ 4.5 H, 1.9 T	≈ 2.5 H, 1.1 T
≈ 21.3 H, 8.6 T	≈ 11.7 H, 8.4 T

$\hat{\theta}_A^{(0)} = 0.60$

$\hat{\theta}_B^{(0)} = 0.50$

1

$$\hat{\theta}_A^{(1)} \approx \frac{21.3}{21.3 + 8.6} \approx 0.71$$

$$\hat{\theta}_B^{(1)} \approx \frac{11.7}{11.7 + 8.4} \approx 0.58$$

3

M-step

4

$\hat{\theta}_A^{(10)} \approx 0.80$

$\hat{\theta}_B^{(10)} \approx 0.52$