BLG527E Machine Learning Homework 4

Elif Özcan 504221508

January 1, 2023

1 Write down the expectation maximization steps to find the parameters. Show the derivation of the formulas that find the parameters (E and M steps).

As seen in Table-1 $\theta_A = \frac{24}{30} = 0.8$ and $\theta_B = \frac{9}{20} = 0.45$.

The Expectation-Maximization steps as follows:

- starts with an initial guess of the parameters
- In the E-step, a probability distribution over possible completions is computed using the current parameters.
- . In the M-step, new parameters are determined using the current completions.
- After several repetitions of the E-step and M-step, the algorithm converges.

1.1 E Step

Let the biases of coin A and B be 0.4 and 0.7 (occurence of heads). For the toss "HHHHHHHHTT" probability distribution for A is:

$$P(E \mid \pi_A) = P(HHHHHHHHTT \mid A) = \frac{10!}{8!2!} * 0.4^8 * 0.6^2$$

and for B is:

$$P(E \mid \pi_B) = P(HHHHHHHHTT \mid B) = \frac{10!}{8!2!} * 0.7^8 * 0.3^2$$

With Bayes theorem we can write the following equation.

$$P(\pi_A \mid E) = \frac{P(E \mid \pi_A)P(\pi_A)}{P(E \mid \pi_A)P(\pi_A) + P(E \mid \pi_B)P(\pi_B)}$$

For the $P(\pi_A)$ and $P(\pi_B)$ we that they are equal to 0.5 so we can write;

$$P(\pi_A \mid E) = \frac{P(E \mid \pi_A)}{P(E \mid \pi_A) + P(E \mid \pi_B)}$$

coin	flips	# coin A heads	# coin B heads
B	HTTTHHTHTH	0	5
A	HHHHTHHHHH	9	0
A	HTHHHHHTHH	8	0
B	HTHTTTHHTT	0	4
A	THHHTHHHTH	7	0

$$P(\pi_A \mid E) = \frac{frac10!8!2! * 0.4^8 * 0.6^2}{frac10!8!2! * 0.4^8 * 0.6^2 + \frac{10!}{8!2!} * 0.7^8 * 0.3^2}$$

$$P(\pi_A \mid E) = \frac{0.4^8 * 0.6^2}{0.4^8 * 0.6^2 + 0.7^8 * 0.3^2} = 0.0435$$

similarly for the coin B:

$$P(\pi_B \mid E) = \frac{0.7^8 * 0.3^2}{0.4^8 * 0.6^2 + 0.7^8 * 0.3^2} = 0.0956$$

flips	**probability it was coin A**	**probability it was coin B**	**# heads in A**	**# heads in B**
HTTTHHTHTH	0.45	0.55	2.2	2.8
HHHHTHHHHH	0.8	0.2	7.2	1.8
HTHHHHHTHH	0.73	0.27	5.9	2.1
HTHTTTHHTT	0.35	0.65	1.4	2.6
THHHTHHHTH	0.65	0.35	4.5	2.5

1.2 M Step

$$\theta_A^1 = \frac{2.2 + 7.2 + 5.9 + 1.4 + 4.5}{10 * (0.45 + 0.8 + *.73 + 0.35 + 0.65)} = 0.71$$

$$\theta_B^1 = \frac{2.8 + 1.8 + 2.1 + 2.6 + 2.5}{10 * (0.55 + 0.2 + 0.27 + 0.65 + 0.35)} = 0.58$$

1.3 Derivations

$$\theta_{n+1} = argmax_\theta\{l(\theta \mid \theta_n\}$$

$$\theta_{n+1} = argmax_\theta\{L(\theta_n) + \sum_z p(z \mid X, \theta_n) ln \frac{p(X \mid z, \theta)p(z \mid \theta)}{p(X \mid \theta_n)p(z \mid X, \theta_n)}\}$$

$$\theta_{n+1} = argmax_\theta\{\sum_z p(z \mid X, \theta_n) lnp(X \mid z, \theta)p(z \mid \theta)\}$$

$$\theta_{n+1} = argmax_\theta\{\sum_z p(z \mid X, \theta_n) ln \frac{p(X, z, \theta)p(z, \theta)}{p(z, \theta)p(\theta}\}$$

$$\theta_{n+1} = argmax_\theta\{\sum_z p(z \mid X, \theta_n) lnp(X, z \mid \theta)\}$$

$$\theta_{n+1} = argmax_\theta\{E_{Z|X,\theta_n}\{p(X, z \mid \theta)\}\}$$