

1 Information gain calculation:

For our problem the information gain is defined as:

$$EIG(\Phi) = \int P(x|\Phi) [H[P(\theta)] - H[P(\theta|x, \Phi)]] dx \quad (1)$$

to compute that we need to sample from $P(x|\Phi)$ and know $P(\theta|x, \Phi)$.

- We don't have $P(x|\Phi)$ because the data of the experiment is conditioned on θ as well, consequently we can only sample from the distribution $P(x|\theta, \Phi)$. But we can calculate it using the following:

$$P(x|\Phi) = \int P(x, \theta|\Phi) d\theta = \int P(x|\theta, \Phi) P(\theta|\Phi) d\theta = E[P(x|\theta, \Phi)] \quad (2)$$

Where the last equality is the expected value of $P(x|\theta, \Phi)$ under the distribution $P(\theta|\Phi)$. Discretising the integral:

$$P(x|\Phi) = \sum_{i=1}^n P(x|\theta_i, \Phi) P(\theta_i|\Phi) = \frac{1}{n} \sum_{i=1}^n P(x|\theta_i, \Phi) \quad (3)$$

Where in the last equality we are assuming to have an uniform distribution $P(\theta_i|\Phi) = 1/n$. We use the black box to generate the distributions $P(x|\theta_i, \Phi)$ for n values of θ :

$$\begin{aligned} \theta_1 &\rightarrow P(x|\theta_1, \Phi) \\ \theta_2 &\rightarrow P(x|\theta_2, \Phi) \\ &\dots \\ \theta_n &\rightarrow P(x|\theta_n, \Phi) \end{aligned} \quad (4)$$

each of these distribution is obtained from a histogram of the sampled data. Using these we can compute $P(x|\Phi)$.

Since $P(x|\Phi)$ is now discretized because of the histogram, we sample from a multinomial with the parameters, x_i 's given by the bin centers and the probability computed from the normalized histogram.

This is done from lines 92 to 107 in the code.

- To calculate the posterior $P(\theta|x_j, \Phi)$ where x_j is sampled using the method described above, we can use Bayes theorem:

$$P(\theta|x_j, \Phi) = \frac{P(x_j|\theta, \Phi) P(\theta|\Phi)}{P(x_j|\Phi)} \quad (5)$$

- After doing this, we average out $-H[P(\theta|x_j, \Phi)]$ across all x_j 's to calculate the Expected Information Gain.