Lehrstuhl für STEUERUNGS- UND REGELUNGSTECHNIK Technische Universität München Prof. Dongheui Lee	**MACHINE LEARNING IN ROBOTICS** Exercises 6: Maximum Likelihood

Exercise 1

Show that the Maximum a posteriori (MAP) estimate becomes Maximum likelihood (ML) estimate if we assume uniform prior distribution for the parameters θ.

Solution:

The MAP estimate is given by

$\arg\max_\theta P(\theta|x^{(1)}, x^{(1)}, \ldots, x^{(n)}) = \arg\max_\theta \frac{P(x^{(1)}, x^{(2)}, \ldots, x^{(n)}|\theta)P(\theta)}{P(x^{(1)}, x^{(2)}, \ldots, x^{(n)})}$

Since denominator is independent of θ, it can be ignored.

$\arg\max_\theta P(\theta|x^{(1)}, x^{(2)}, \ldots, x^{(n)}) \propto \arg\max_\theta P(x^{(1)}, x^{(2)}, \ldots, x^{(n)}|\theta)P(\theta)$

Now if the prior $P(\theta)$ is uniformly distributed then it can be replaced with a constant and the maximization is reduced to

$\arg\max_\theta P(\theta|x^{(1)}, x^{(2)}, \ldots, x^{(n)}) \propto \arg\max_\theta P(x^{(1)}, x^{(2)}, \ldots, x^{(n)}|\theta)$

Now as the smaples are independent and identically distributed

$\arg\max_\theta P(\theta|x^{(1)}, x^{(2)}, \ldots, x^{(n)}) \propto \arg\max_\theta P(x^{(1)}|\theta)P(x^{(2)}|\theta)\ldots P(x^{(n)}|\theta) = \arg\max_\theta \prod_{i=1}^{n} P(x^{(i)}|\theta)$

which is the well known maximum likelihood approach for parameters estimation.

Exercise 2

Suppose that there is a box with three coins and the probability of head for each coin is $P(H|c_1) = \frac{1}{3}$, $P(H|c_2) = \frac{1}{2}$ and $P(H|c_3) = \frac{2}{3}$. One coin was picked at random and tossed 100 times. The result is 49 heads and 51 tails. Predict the coin.

Solution:

$P(H|c_1) = \frac{1}{3} \Rightarrow P(T|c_1) = 1 - \frac{1}{3} = \frac{2}{3}$
$P(H|c_2) = \frac{1}{2} \Rightarrow P(T|c_2) = 1 - \frac{1}{2} = \frac{1}{2}$
$P(H|c_3) = \frac{2}{3} \Rightarrow P(T|c_3) = 1 - \frac{2}{3} = \frac{1}{3}$

The samples can be regarded to have i.i.d distribution:

$L(\theta) = \prod_{i=1}^{n} P(x^{(i)}|\theta)$

For numerical stability it is better to use logarithm

$l(\theta) = \log L(\theta)$

Now the likelihood value for each coin is:

Likelihood value for $c_1 = \sum_{i=1}^{n} \log P(x^{(i)}|c_1) = 49 \log \frac{1}{3} + 51 \log \frac{2}{3} = -74.5107$

Likelihood value for $c_2 = \sum_{i=1}^{n} \log P(x^{(i)}|c_2) = 49 \log \frac{1}{2} + 51 \log \frac{1}{2} = -69.3147$

Likelihood value for $c_3 = \sum_{i=1}^{n} \log P(x^{(i)}|c_3) = 49 \log \frac{2}{3} + 51 \log \frac{1}{3} = -75.8970$

So our guess is c_2 as it yields maximum value for likelihood.

Exercise 3

The Poisson distribution is a discrete probability distribution that expresses the probability of a given number of events occurring in a fixed interval of time. The probability mass function is given by :
$P(X = k|\mu) = \frac{\mu^k e^{-\mu}}{k!}$ where $\mu > 0$.

Now consider a factory in which a group of industrial robots are used for manufacturing the automobile parts. Sometimes the parts are found to be defective. The number of defective parts produced in n different months are given by $X = x^{(1)}, x^{(2)}, \ldots, x^{(n)}$, which are assumed to be i.i.d. poisson random variables.

a) Use the samples to get a maximum likelihood estimate of μ.
b) For $\mu = 10$, find the probability of producing more than 4 defective parts in a month.

Solution:

a) $L(\theta) = \prod_{i=1}^{n} \frac{e^{-\mu}\mu^{x^{(i)}}}{x^{(i)}!} = \frac{e^{-n\mu}\mu^{\sum_{i=1}^{n} x^{(i)}}}{\prod_{i=1}^{n} x^{(i)}!}$

$l(\theta) = -n\mu + \sum_{i=1}^{n} x^{(i)} \log \mu - \log \prod_{i=1}^{n} x^{(i)}!$

Evaluating the gradient of $l(\mu)$ to zero we get

$$-n + \frac{\sum_{i=1}^{n} x^{(i)}}{\mu} = 0 \tag{1}$$
$$\mu = \frac{\sum_{i=1}^{n} x^{(i)}}{n} \tag{2}$$
$$\mu = \bar{x} \tag{3}$$

b)

$$\begin{aligned}
P(X > 4) &= 1 - P(X \leq 4) \\
&= 1 - \sum_{k=0}^{4} \frac{\mu^k e^{-\mu}}{k!} \\
&= 1 - \frac{10^0 e^{-10}}{0!} - \frac{10^1 e^{-10}}{1!} - \frac{10^2 e^{-10}}{2!} - \frac{10^3 e^{-10}}{3!} - \frac{10^4 e^{-10}}{4!} \\
&= 1 - 0.0293 = 0.9707
\end{aligned}$$