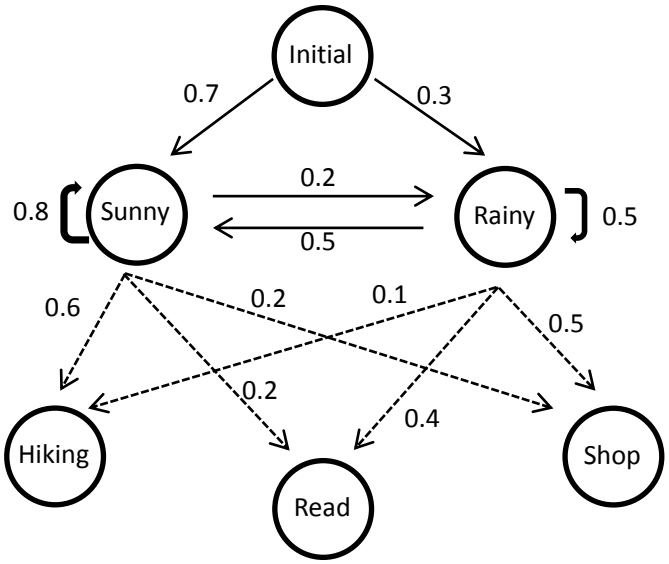
Hong Kong University of Science and Technology
COMP4211: Machine Learning
Spring 2013

Assignment 1

Due: 29/4/2013, Monday, in class

1. (60 points)
   Bob is interested in three activities: hiking, reading and shopping. His choice of what to do depends on the weather condition of the given day. Now, there are two states of the weather: "Sunny" and "Rainy". Below is the transition network and the output distribution of given state.

   It is now known that Bob's activities for these three days are: **hiking, reading and shopping** sequentially.



   1) Derive the most likely weather condition sequence for these three days using Viterbi algorithm.

   2) Using forward probability, derive the probability that Bob's activities for these three days are: **hiking, reading and shopping**.

   3) Using backward probability, derive the probability that Bob's activities for these three days are: **hiking, reading and shopping**.

   4) Use the EM algorithm (forward-backward algorithm) to update the parameters of the model once.

2. (40 points)

Considering the multinomial distribution, its probability mass function is given by:

$$\frac{(\sum_i \alpha_i)!}{\prod_i \alpha_i!} \prod x_i^{\alpha_i}$$

where the count of each category is $\alpha_1, \dots, \alpha_k$ and the probability for each of the k category is $x_1, \dots, x_k$.

We are rolling a dice with 6 possible outcomes, given that it is under multinomial distribution.

1) Suppose we now have an observation as follow:

Outcome	**1**	**2**	**3**	**4**	**5**	**6**
Number of occurrence	1	2	5	2	5	2

and the probability of each outcome is:

Outcome	**1**	**2**	**3**	**4**	**5**	**6**
Probability x_i:	1/12	1/6	1/6	1/6	1/4	1/6

What is probability of the occurrence of the given observation?

2) Suppose we have the following observation **α**:

Outcome	**1**	**2**	**3**	**4**	**5**	**6**
Pseudo count of occurrence α_i:	3	2	6	4	6	3

and we have three probability possibilities for each outcome (let's denote the parameters set as θ):
First case $\boldsymbol{\theta_1}$:

Outcome	**1**	**2**	**3**	**4**	**5**	**6**
Probability x_i:	1/12	1/6	1/6	1/6	1/4	1/6

Second case $\boldsymbol{\theta_2}$:

Outcome	**1**	**2**	**3**	**4**	**5**	**6**
Probability x_i:	1/3	1/12	1/12	1/3	1/12	1/12

Third case $\boldsymbol{\theta_3}$:

Outcome	**1**	**2**	**3**	**4**	**5**	**6**
Probability x_i:	1/16	1/16	1/4	1/4	1/4	1/16

i. Using Dirichlet distribution, what is the value of $p(\boldsymbol{\theta_1}), p(\boldsymbol{\theta_2})$ and $p(\boldsymbol{\theta_3})$, that is the probability density function values of the given parameters of the distribution?

(The probability density function of Dirichlet distribution is:

$$p(\boldsymbol{\theta}) = f(x_1, \dots, x_k |\ \alpha_1, \dots, \alpha_k) = \frac{\Gamma(\sum_i \alpha_i)}{\prod_i \Gamma(\alpha_i)} \prod x_i^{\alpha_i - 1}$$

where the probability for each of the k category is $x_1, \dots, x_k$ under the restriction that $\sum_i x_i = 1$ and the **pseudo count** of each category are $\alpha_1, \dots, \alpha_k$.

You should notice that, in Dirichlet distribution, $\boldsymbol{\alpha_i}$ **is the pseudo count** and $\boldsymbol{\alpha_i - 1}$ **is the actual number of occurrence**.)

ii. By the observation in part (i), we have got $p(\boldsymbol{\theta_1}), p(\boldsymbol{\theta_2})$ and $p(\boldsymbol{\theta_3})$ for each case.

Suppose we now have **another** observation $\boldsymbol{\alpha'}$:

Outcome	**1**	**2**	**3**	**4**	**5**	**6**
Pseudo count of occurrence $\boldsymbol{\alpha'_i}$:	2	2	5	3	6	3

Calculate the $p(\boldsymbol{x'}|\boldsymbol{\theta_1}), p(\boldsymbol{x'}|\boldsymbol{\theta_2})$ *and* $p(\boldsymbol{x'}|\boldsymbol{\theta_3})$ for this observation. Using MAP (maximum a posterior) estimation:

$$\boldsymbol{\theta'} = argmax_{\theta}\ [p(\boldsymbol{x'}|\boldsymbol{\theta})p(\boldsymbol{\theta})]$$

find out which case is the most likely one under this newly given observation $\boldsymbol{\alpha'}$.