Xi'an Jiaotong-Liverpool University

西交利物浦大学

XJTLU Entrepreneur College (Taicang) Cover Sheet

Module code and Title	**DTS201TC Pattern Recognition**	
School Title	**School of AI and Advanced Computing**	
Assignment Title	**Final project**	
Submission Deadline	**23:59, 31st Dec.**	
Final Word Count		
If you agree to let the university use your work anonymously for teaching and learning purposes, please type **"yes"** here.		

I certify that I have read and understood the University's Policy for dealing with Plagiarism, Collusion and the Fabrication of Data (available on Learning Mall Online). With reference to this policy I certify that:

- My work does not contain any instances of plagiarism and/or collusion.
  My work does not contain any fabricated data.

By uploading my assignment onto Learning Mall Online, I formally declare that all of the above information is true to the best of my knowledge and belief.

Scoring – For Tutor Use	
Student ID	1931097

Stage of Marking	**Marker Code**	**Learning Outcomes Achieved (F/P/M/D) (please modify as appropriate)**			**Final Score**
		A	**B**	**C**	
1st Marker – red pen					
Moderation – green pen	**IM Initials**	The original mark has been accepted by the moderator (please circle as appropriate):			Y / N
		Data entry and score calculation have been checked by another tutor (please circle):			Y
2nd Marker if needed – green pen					
For Academic Office Use		**Possible Academic Infringement (please tick as appropriate)**			
Date Received	**Days late**	**Late Penalty**	☐ **Category A**	Total Academic Infringement Penalty (A,B, C, D, E, Please modify where necessary) _____________________	
			☐ **Category B**		
			☐ **Category C**		
			☐ **Category D**		
			☐ **Category E**		

DTS201TC Classification Demonstration

Project (Individual)

1 Mathematical problems [40 marks]

Derive the Maximum Likelihood Estimate.

1.1 [20 marks]

Let $x_1, x_2, ..., x_N$ be vectors stemmed from a normal distribution with known covariance matrix and unknown mean, that is

$$p(x_k; \mu) = \frac{1}{(2\pi)^{D/2}|\Sigma|^{1/2}} exp(-\frac{1}{2}(x_k - \mu)^T \Sigma^{-1}(x_k - \mu)) \tag{1}$$

where D is the dimension of vector x_k $(k = 1, ..., N)$.

***TASK 1:* Derive the ML estimate of the mean μ.**
Solution:

According to the definition of log-likelihood function:

$$L(\mu) = \ln \prod_{k=1}^{N} p(x_k; \mu) \tag{2}$$

Substituting the given distribution (1) into equation (2) and simplifying it leads to

$$L(\mu) = -\frac{DN}{2} \ln 2\pi - \frac{N}{2} \ln|\Sigma| - \frac{1}{2} \sum_{k=1}^{N} (x_k - \mu)^T \Sigma^{-1}(x_k - \mu) \tag{3}$$

Gradient of equation (3) with respect to μ is equivalent to

$$\frac{\partial L(\mu)}{\partial \mu} = \frac{\partial(-\frac{1}{2}\sum_{k=1}^{N}(x_k - \mu)^T \Sigma^{-1}(x_k - \mu))}{\partial \mu} \tag{4}$$

To simplify equation (4), multiply both the numerator and denominator by $\partial(x_k - \mu)$,

$$\begin{aligned} \frac{\partial L(\mu)}{\partial \mu} &= \frac{\partial(-\frac{1}{2}\sum_{k=1}^{N}(x_k - \mu)^T \Sigma^{-1}(x_k - \mu))}{\partial(x_k - \mu)} \frac{\partial(x_k - \mu)}{\partial \mu} \\ &= (-\frac{1}{2} \sum_{k=1}^{N} (x_k - \mu)(x_k - \mu)^T \Sigma^{-1})(-1) \\ &= \sum_{k=1}^{N} \Sigma^{-1}(x_k - \mu) \end{aligned} \tag{5}$$

Setting the partial derivative to be 0, we have

$$\frac{\partial L(\mu)}{\partial \mu} = \sum_{k=1}^{N} \Sigma^{-1}(x_k - \mu) = 0 \tag{6}$$

Therefore

$$\mu = \frac{1}{N} \sum_{k=1}^{N} x_k \tag{7}$$

1.2 [20 marks]

Let $x_1, x_2, ..., x_N$ be vectors stemmed from a normal distribution with unknown mean μ and unknown convariance matrix Σ, that is

$$p(x_k; \mu, \Sigma) = \frac{1}{(2\pi)^{D/2}|\Sigma|^{1/2}} exp(-\frac{1}{2}(x_k - \mu)^T \Sigma^{-1}(x_k - \mu)) \tag{8}$$

where D is the dimension of vector x_k $(k = 1, ..., N)$.

***TASK 2:* Derive the ML Estimate of μ and Σ.**
Solution:

Same as 1.1, substituting the given distribution (9) into equation (2) and simplifying it leads to

$$L(\mu, \Sigma) = -\frac{DN}{2} \ln 2\pi - \frac{N}{2} \ln|\Sigma| - \frac{1}{2} \sum_{k=1}^{N} (x_k - \mu)^T \Sigma^{-1}(x_k - \mu) \tag{9}$$

Getting the partial derivative, then simplifying

$$\begin{aligned} \frac{\partial L(\mu, \Sigma)}{\partial \mu} &= \frac{\partial(-\frac{1}{2}\sum_{k=1}^{N}(x_k - \mu)^T \Sigma^{-1}(x_k - \mu))}{\partial \mu} \\ &= \sum_{k=1}^{N} \Sigma^{-1}(x_k - \mu) \end{aligned} \tag{10}$$

$$\begin{aligned} \frac{\partial L(\mu, \Sigma)}{\partial \Sigma} &= \frac{\partial(-\frac{N}{2}\ln|\Sigma|)}{\partial \Sigma} + \frac{\partial(-\frac{1}{2}\sum_{k=1}^{N}(x_k - \mu)^T \Sigma^{-1}(x_k - \mu))}{\partial \Sigma} \\ &= -\frac{N}{2}\Sigma^{-1} + \frac{-\frac{1}{2}\sum_{k=1}^{N}(x_k - \mu)^T \mathrm{d}\Sigma^{-1}(x_k - \mu)}{\mathrm{d}\Sigma} \end{aligned} \tag{11}$$

According to the rule, we have

$$0 = \mathrm{d}I = \mathrm{d}(\Sigma\Sigma^{-1}) = \mathrm{d}\Sigma\Sigma^{-1} + \Sigma\mathrm{d}\Sigma\Sigma^{-1}$$

Thus

$$\mathrm{d}\Sigma^{-1} = -\Sigma^{-1}\mathrm{d}\Sigma^{-1} \tag{12}$$

Substitute the equation (13) into the equation (12)

$$\begin{aligned} \frac{\partial L(\mu, \Sigma)}{\partial \Sigma} &= -\frac{N}{2}\Sigma^{-1} + \frac{\frac{1}{2}\sum_{k=1}^{N}(x_k - \mu)^T - \Sigma^{-1}\mathrm{d}\Sigma\Sigma^{-1}(x_k - \mu)}{\mathrm{d}\Sigma} \\ &= -\frac{N}{2}\Sigma^{-1} + \frac{1}{2}\sum_{k=1}^{N}(x_k - \mu)^T - \Sigma^{-1}\Sigma^{-1}(x_k - \mu) \end{aligned} \tag{13}$$

Setting the partial derivative (11) (14) to be 0, we have

$$\frac{\partial L(\mu, \Sigma)}{\partial \mu} = \sum_{k=1}^{N} \Sigma^{-1}(x_k - \mu) = 0 \tag{14}$$

$$\frac{\partial L(\mu, \Sigma)}{\partial \Sigma} = -\frac{N}{2}\Sigma^{-1} + \frac{1}{2}\sum_{k=1}^{N}(x_k - \mu)^T - \Sigma^{-1}\Sigma^{-1}(x_k - \mu) = 0 \tag{15}$$

Therefore,

$$\mu = \frac{1}{N}\sum_{k=1}^{N} x_k \tag{16}$$

$$\Sigma = \frac{1}{N}\sum_{k=1}^{N}(x_k - \mu)(x_k - \mu)^T = \frac{1}{N}\sum_{k=1}^{N}(x_k - \bar{x})(x_k - \bar{x})^T \tag{17}$$

2 Practical problems [60 marks]

This assignment is designed to deal with MSRC-12 Kinect gesture data set of Microsoft Research Cambridge. The dataset contains lots of sequences of skeletal body movements recorded from a kinect device. However, this assignment only consider a small fraction of it, which consists of static body positions.

Assignment: classify the body positions with Bayes model.
Data: MSRC-12 Kinect gesture dataset of Microsoft Research Cambridge (provided on LMO together with this sheet)

- The dataset consists of 2045 instances of body positions with 4 categories of positions. “arms lifted”, “right arm extended to one side”, “crouched” and “right arm extended to the front”. These classes are represented with distinctive numbers.
- The body positions are encoded with a 20×3 matrix, in which the row is the position in space (x,y,z) of each of the 20 joints. Each variable is modeled with a Gaussian distribution.
- Assume the 60 variables that define the body position are considered independent given the class.

Formulation: To be specific, you are expected to use Naive Bayes model. In Naive Bayes model, each of the 60 variables is considered independent given the class. Each variable is modeled with a Gaussian distribution. The training process for the model is to estimate the values for the mean and variance for each variable and class with MLE.

$$p(x_i|C) = Normal(\mu_{x_i}; \sigma^2_{x_i})$$
$$p(y_i|C) = Normal(\mu_{y_i}; \sigma^2_{y_i})$$
$$p(z_i|C) = Normal(\mu_{z_i}; \sigma^2_{z_i})$$

Where $p(i)$ is the probability density function of i-th joint,

$$P(C = k|sample) \propto P(C = k) \prod_{i=1}^{N} p(x_i, y_i, z_i|C = k)$$
$$p(x_i, y_i, z_i|C = k) = p(x_i|C = k) \times p(y_i|C = k) \times p(z_i|C = k)$$

Where, N is the number of samples, k indicates k-th label.

2.1 TASK 3:[20 marks]

Implement the function to estimate the parameters of the Gaussian distributions using MLE.

function: ***fit_model***

- Input: a vector which is the observation for a given variable
- Output: the mean and the standard deviation for these observations.

2.2 TASK 4: [20 marks]

Implement a function to build a model composed of priors, and model parameters.

function :***learn_model***

- Input: the dataset and the labels
- Output: compute the parameters from the dataset to build the model

2.3 TASK 5: [20 marks]

Implement the classification function
function :***classify_samples***

- Input: a set of instances that have the same format as the dataset given in *learn_model*
- Output: posterior probability for each instance belonging to each class

Marking Scheme for Task 3-5:

- You can choose to write Pseudocode in stead of implementing functions by coding.
- Write Pseudocode will get 2-10 marks for each function.
- Codes of each function can run properly, and can output accuracy on the test data, the accuracy is incorrect: incorrectly implemented function will get 5-15 marks accordingly.
- Codes of each function can run properly, and can output accuracy on the test data, the accuracy is correct : 60 marks in total.