CSCI567 2014 Homework Assignment 5

Student Name Yi Duan

USC ID 9966845562

1. Principal Component Analysis
   1. Deriving PCA in terms of minimum reconstruction error
2. Set

Take derivation of J respect to ,

Then we get

1. Substitute with and we get

Impose the constraint and we minimize

Take derivation of respect to ***U***

We get

Where is the covariance matrix. And we find ***U*** is the eigenvectors of .

* 1. Projecting a Gaussian distribution

1. Given and , we know . Thus,

And the expectation of z2, E(z2), is equal to , that is,

The entropy of z is

According to and , we get

Thus,

Impose the constraint and we maximize

Take derivation of respect to

Then we get optimal is the eigenvalue of

1. We know maximize from above. Thus, According to , maximize the variance of z.
2. Hidden Markov Models

Thus, the most likely path of hidden states is 2,2,2,2,1,1. While the sequence of most likely states estimated independently is 2,1,2,2,1,1.



Thus, the most likely emitted symbol at the end of sequence e is “A” and “G”.

1. Programming (PCA)
2. The top 5 eigenfaces are as follow:

F:\Dropbox\567\assignment\hw5 - 副本\eigenface1.bmp F:\Dropbox\567\assignment\hw5 - 副本\eigenface2.bmp F:\Dropbox\567\assignment\hw5 - 副本\eigenface3.bmp F:\Dropbox\567\assignment\hw5 - 副本\eigenface4.bmp F:\Dropbox\567\assignment\hw5 - 副本\eigenface5.bmp

1. Optimal parameters

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| SVM | Parameter | d=20 | d=50 | d=100 | d=200 |
| Linear | Log4C | -2 | 2 | 2 | 3 |
| RBF kernel | Log4C | 6 | 6 | 6 | 5 |
| Log4g | -8 | -7 | -8 | -7 |

Average test accuracy

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| SVM | d=20 | d=50 | d=100 | d=200 |
| Linear | 74.4173 | 75.6353 | 75.2281 | 75.7744 |
| RBF kernel | 88.2845 | 90.8709 | 91.2105 | 91.0376 |

1. Programming (HMM)
2. Hmm\_fun

A\_estimate

0.9305 0.0695

0.0349 0.9651

E\_estimate

0.3711 0.1227 0.1290 0.3773

0.0716 0.4465 0.4217 0.0602

hmmtrain

A\_estimate

0.9249 0.0751

0.0844 0.9156

E\_estimate

0.3850 0.1048 0.1165 0.3937

0.0891 0.4386 0.4003 0.0720