COMP5318 Quiz Week 6, Sem 1, 2020

Due No due date **Points** 100 **Questions** 25
Available Sep 28 at 18:10 - Sep 28 at 20:02 about 2 hours
Time Limit 60 Minutes

Instructions

$$A = U \sum V^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 0.04 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

Assume A has the above SVD decomposition and each column of A represents one entity (e.g. customer, image, document).

Questions 1-5 are based on the above decomposition. Answer the following multiple choice questions.

For questions 6-10:
We have learnt that many data types can be represented using a matrix. Also, using the notion of a dot product we can capture the similarity between two entities.

Questions 11-15 are about probability theory and Bayes' rule.

Questions 16-19 are on Python programming.

Questions 20-25 are on KNN, Linear Regression, Gradient descent, Overfiting.

Attempt History

	Attempt	Time	Score
LATEST	Attempt 1	33 minutes	100 out of 100

Correct answers are hidden.

Score for this quiz: **100** out of 100
Submitted Sep 28 at 19:18
This attempt took 33 minutes.

Question 1

4 / 4 pts

Which of the following statements below is INCORRECT?

- ○ Both U and V have the same rank.
- ◉ V is a diagonal matrix.
- ○ U is a symmetric matrix.
- ○ Columns of U and V are basis vectors.
- ○ V is an orthogonal matrix.

Question 2

4 / 4 pts

Which of the following is the most accurate statement?

- ○ There is exactly one concept.
- ◉ There are 2 concepts, and one of them has more impact than the other.
- ○ There are more than 3 "concepts" (a.k.a, latent dimensions) of A.

- There are 2 concepts with equivalent impact.

Question 3

4 / 4 pts

The SVD decomposition shows that the rank of the matrix A, denoted by rank(A), is:

- 3
- (•) The same rank as both U and V.
- 1
- rank (U + V) + 2
- rank(U - V) - 3.
- rank(UV) + 3.

Question 4

4 / 4 pts

Given S, set of singular values of the given matrix A, the best rank 2 approximation of A is obtained by taking?

-

The product of the first column of U, the first element of S, and the first row of V^T.

- [x] Matrix A itself.
- [ ] The product of U and S.
- [ ] The product of the second column of U, the second element of S and the second row of V^T.
- [ ] The product of U and V^T.
- [ ] The second column of V^T.

Question 5

4 / 4 pts

Which of the following is the most accurate statement?

- [ ] The data (represented by A) is well spread out in all directions.
- [x] The data is spread out mostly in the direction of a vector scaled by the largest singular value of A.
- [ ] The data is spread out mostly in the direction of a vector scaled by the smallest singular value of A.
- [ ] The data is equally spread out in the two dimensional plane.

Question 6

4 / 4 pts

Suppose x = (1, 2, -3, 1) and y = (1, 1, 2, 2). Then the dot product and cosine similarity of x and y are respectively given by?

- 1.2 and 0.1
- -1 and -0.105
- -8 and -0.2
- (•) -1 and -0.081
- 12 and -1

Question 7

4 / 4 pts

The lengths of vector x and y are:

- 3.87 and 2.64
- 3.87 and 2.45
- (•) 3.87 and 3.16
- 3.75 and 2.52
- 3.11 and 2.22

- 2.65 and 2.24

Question 8

4 / 4 pts

Which of the following are orthogonal vectors?

- (3, 2) and (-2, 2).
- $(\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ and $(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}})$.
- (1, 1) and (-2, 1).
- (1, -1, 1) and (-1, 1, 2).
- (1, -1) and (-1, 1).

Question 9

4 / 4 pts

Which of the following pair of vectors are most similar in cosine similarity to each other?

- (1, 1, 1, 1) and (3, 3, 3, 4).
- (1.2, 1.2, 1.2, 1.2) and (2, 2, 2, 2).
- (1, -1, 1, -1) and (1, 1, 1, 1).

(1, 1, 1, 1) and (1, 1, 1, 2).

(1, 2, 3, 1) and (3, 2, 1, 2).

Question 10

4 / 4 pts

Given a SVD decomposition of a matrix $B = U\Sigma V^T$ of size (m,n). Then which is the most accurate statement?

U is a (m,n) column-orthogonal matrix, $\sum$ is diagonal $r \times r$ matrix, and V is a (n,r) column-orthogonal matrix

U is a (m,n) column-orthogonal matrix, $\sum$ is diagonal $m \times n$ matrix, and V is a (n,r) column-orthogonal matrix

U is a (m,r) column-orthogonal matrix, $\sum$ is diagonal $r \times r$ matrix, and V is a (n,m) column-orthogonal matrix

U is a (m,r) column-orthogonal matrix, $\sum$ is diagonal $r \times r$ matrix, and V^T is a (r,n) row-orthogonal matrix

Question 11

4 / 4 pts

There is a black box with 8 balls in it, including 3 green balls, 3 red balls, and remaining white balls. When randomly pick one ball from the black box, what is the probability that the picked ball is white?

- 0.47
- 0.25
- 0.12
- None of the other answers.
- 0.61

Question 12

4 / 4 pts

When two consecutive picking actions are performed and the picked ball will be placed back to the blackbox before performing a new picking, what is the probability that first picked ball is red and the second picked ball is green?

- 0.30
- None of the other answers.
- 0.24
- 0.36
- 0.16

Question 13

4 / 4 pts

Suppose P(A) = 0.2 and P(B) = 0.2 and P(A|B) = 0.5 then what is P(A U B) ?

- 0.15
- 0.1
- None of the other answers.
- 0.25
- 0.3
- 0.2

Question 14

4 / 4 pts

Suppose P(A) = 0.2 and P(B) = 0.2 and P(A|B) = 0.5 then what is P(B|A) ?

- 0.25
- 0.15
- None of the other answers.

- (•) 0.5
- ( ) 0.2
- ( ) 0.1

Question 15

4 / 4 pts

Suppose P(A) = 0.2 and P(B) = 0.2 and P(A|B) = 0.5 then what is the probability that both event A and B will not occur?

- ( ) 0.5
- ( ) 0.3
- ( ) 0.2
- ( ) 0.4
- (•) 0.7

Question 16

4 / 4 pts

In order to add the number 5 to the end of the list s = [1, 2, 3, 4], what function should you use?

- A. s.append(5)
- B. s[-1] = 5
- C. s.insert(s,5)
- D. s = s+[5]

- ○ A, B, C, D
- ○ A, C, D
- ◉ A, D
- ○ B, C, D

Question 17

4 / 4 pts

Assume that 2 arrays A and F are defined by numpy as follows:

import numpy as np

A = np.array([1, 2, 3])

F = np.array([[4, 2, 11, 3]])

Which code is correct to calculate $A^T . F$?

- ○ A[:,np.newaxis].T.dot(F)

- A.T.F
- A.T.dot(F)
- A[np.newaxis, :].T.dot(F)

Question 18

4 / 4 pts

What is the result of this python script?

```
d = {'apple':10, 'pear':20,'orange':10, 'berry': 'sold out'}
d['banana'] = 30
values = set([d[key] for key in d.keys()])
print(values)
```

- {'apple','pear','berry','banana'}
- {'sold out', 10, 20, 30}
- {'apple','pear','berry'}
- {10, 20, 10, 30, 'sold out'}
- (10, 20, 10, 'sold out', 30)
- (10, 20, 10, 30, 'sold out')

Question 19

4 / 4 pts

What is the matrix y close to?

```
import numpy as np
x = np.random.random((5,5))
x = x + x.T
u, v = np.linalg.eig(x)
y = v.T.dot(v)
```

- dot(u.T, u)
- x
- cannot be determined
- (•) eye(5)
- zeros((5,5))

Question 20

4 / 4 pts

Which of the following evaluation metrics is the best to evaluate a linear regression model?

- (•) Mean-Squared-Error
- Confusion Matrix
- AUC-ROC

Accuracy

Question 21

4 / 4 pts

Suppose we use gradient descent to try to minimize function RSS(w). Which statement about Gradient Descent is correct?

- (•) If the learning rate is too small, then gradient descent may take a very long time to converge.
- ( ) Setting the learning rate to be very large is not harmful, and can speed up the convergence of gradient descent.
- ( ) Gradient Descent can only be used for Regression Problems
- ( ) Even if the learning rate is very large, every iteration of gradient descent will decrease the value of f(w)

Question 22

4 / 4 pts

In polynomial regression, suppose we have a nonlinear mapping such as:

$$\phi(\boldsymbol{x}) : \boldsymbol{x} \in \mathbb{R}^D \rightarrow \boldsymbol{z} \in \mathbb{R}^M$$

Which statement is incorrect to prevent overfitting for polynomial regression?

- (•) Increase the value of M
- ( ) Decrease the value of M
- ( ) Train with more data
- ( ) Using regularization

Question 23

4 / 4 pts

Which statement about Naive Bayes is not correct?

- ( ) Naive Bayes classifier could deal with continuous features.
- (•) Naive Bayes classifier assumes a strong dependence assumption between features.
- ( )

Naive Bayes classifier is robust to isolated noise points and irrelevant features.

Naive Bayes classifier is a probabilistic model.

Question 24

4 / 4 pts

Instance	P(+|A)	True Class
1	0.8	+
2	0.3	+
3	0.4	-
4	0.1	-

Given 4 instances with true class and posterior probability produced by some classifier, where P(+|A) is the probability of A is classified as positive for each instance A. What is the false positive rate when the threshold is 0.3 (the classifier would classify instances with P(+|A)>=0.3 as positive)? Hint: FPR = FP/(FP + TN).

0.33

0.67

0.5

0

1

Question 25

4 / 4 pts

Which statement about K-nearest neighbour is not correct?

- (•) When K = 1, the KNN classifier may be faced with underfitting problem.
- ( ) KNN classifier is a non-linear model.
- ( ) KNN is a non-parametric model.
- ( ) Euclidean distance is a common metric to measure the distance between data points in KNN classifier.

Quiz Score: **100** out of 100