CSE 152 Mid-Term Examination

University of California San Diego

Oct 29, 2019

Name: ______________________________

UCSD ID: ______________________________@ucsd.edu

Question	**Points**	**Points Earned**
1	10	
2	10	
3	10	
4	10	
5	10	
6	10	
7	10	
8	10	
9	10	
Total	90	

Instructions:

1. This examination contains 12 pages, including this page.
2. You have **one (1) hour** to complete the examination. As a courtesy to your classmates, we ask that you not leave during the last fifteen minutes.
3. Write your answers in this booklet. We scan this into Gradescope, so **please try to avoid writing on the backs of pages**. If you must do so, please indicate **very** clearly on the front of the page that you have written on the back of the page.
4. You may use two (2) double-sided 8.5" $\times$ 11" pages with notes that you have prepared. You may not use any other resources, including lecture notes, books, other students or other engineers.
5. You may use a calculator. You may not share a calculator with anyone. If you didn't bring a calculator, you may use your phone, **but** you must put it on **flight mode** and clear all visible notifications **before** the examination starts, and you must not open any applications other than the calculator and a timer.

Question 1: Basic Linear Algebra

$$A = \begin{bmatrix} 1 & 0 & 1 \\ \lambda & 1 & 1 \\ 1 & 1 & \lambda \end{bmatrix}$$

(a) Calculate the rank of A when $\lambda = 1$. Is A invertible? [2 pts]

(b) What's the null space of A when $\lambda = 1$? What's the dimension of its null space? [3 pts]

(c) Give the condition of λ when A is a full rank matrix. What is the dimension of A's null space for such λ's? [5 pts]

Question 2: Least Squares Problem

(a) For an over-determined linear system $Ax = b$, where A is an $m \times n$ tall matrix ($m > n$) with linearly independent columns (full rank). When would this linear system have no solution? (Hint: write A as matrix of its columns) [2 pts]

(b) Write down the least squares problem formulation [2 pts], and derive the solution [2 pts].

(c) Consider an under-determined linear system, where A is a fat matrix ($m < n$). How many solution(s) does $Ax = b$ have? Select from below (you can make multiple choices if more than one could happen). Briefly explain why. [2 pts]

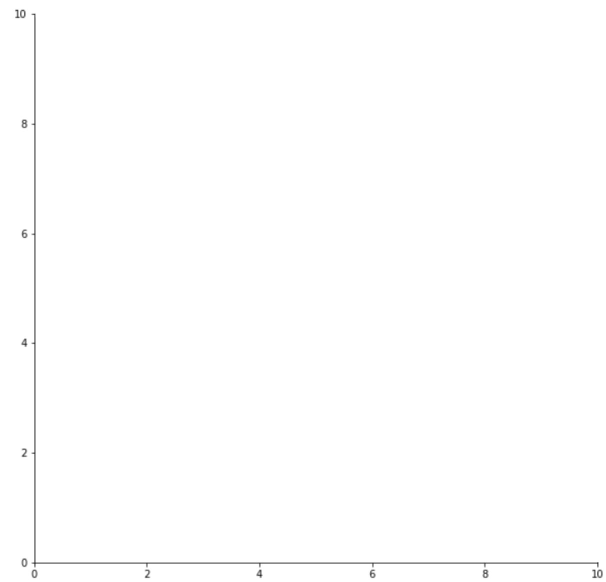
No solution | one solution | infinite solutions

(d) In class, we taught that we may obtain a special solution for this under-determined system by $x = A^{\dagger}b$ using the pseudo inverse, where $A^{\dagger} = A^T(AA^T)^{-1}$. What is the property (geometric interpretation) of this special solution? [2 pts]

Question 3: K-Means

Point	X Coord.	Y Coord.
1	6	2
2	7	1
3	8	3
4	7	2
5	3	7
6	1	9
7	2	7
8	2	5

(a) Plot the points on the graph and draw the circles around the two clusters of data. [2 pts]



b) The first iteration of your k-means algorithm initialized the cluster centers at $C_1 = (2,6)$ and $C_2 = (6,2)$. Calculate the distance between Point 2 and it's cluster center. [3 pts]

c) Using the work you have done in part a), draw the (rough) new locations for clusters C_1 and C_2 for the next iteration of the algorithm. [5 pts]

Question 4: Bag of Visual Words

Creating and implementing a Bag of Words model involves the following steps:

$$\begin{array}{c}\text{Extract Features}\\ \downarrow \\ \text{Learn Visual Vocab}\\ \downarrow \\ \text{Quantize Features}\\ \downarrow \\ \text{Represent Image by Frequency of Visual Words}\end{array}$$

a) Explain what occurs at each stage of the process. [8 pts]

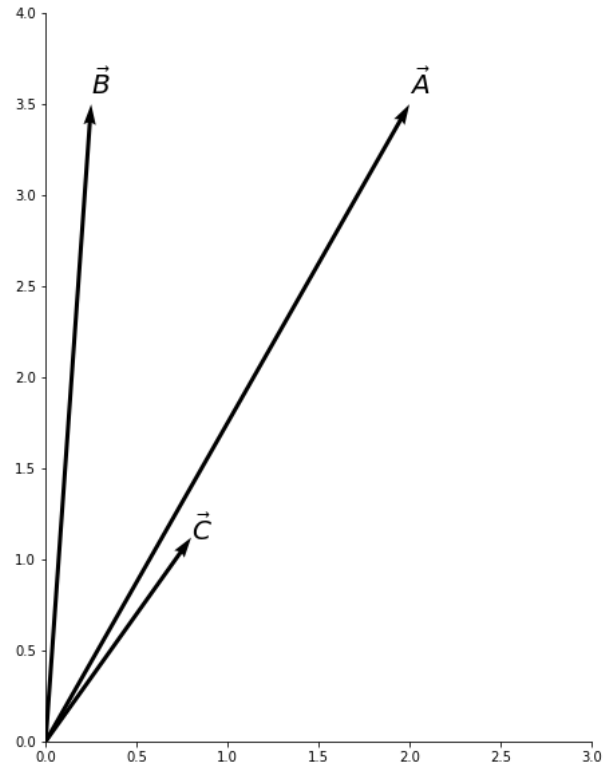
<u>Extract Features:</u>

<u>Learn Visual Vocab:</u>

<u>Quantize Features:</u>

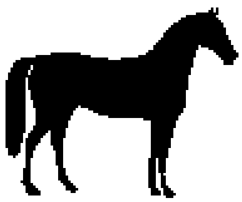
<u>Represent Image by Frequency of Visual Words:</u>

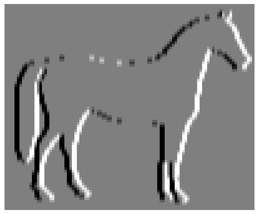
b) We utilize the cosine similarity metric to compare the visual words distributions between images. In the image below, indicate which vector has the higher cosine similarity to $\vec{A}$ by circling the vector letter. [2 pts]



Question 5: Filters

(a) An input image shown below is filtered by some image filters to produce A, B, and C. Briefly explain the effect of each filter. [6 pts]


input


A


B


C

1. $\frac{1}{9}\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{bmatrix}$

2. $\begin{bmatrix}-1 & 0 & 1\\ -2 & 0 & 2\\ -1 & 0 & 1\end{bmatrix}$

3. $\begin{bmatrix}-1 & -2 & -1\\ 0 & 0 & 0\\ 1 & 2 & 1\end{bmatrix}$

Explain.

(b) Given a gray-scale image I with shape (H, W). Let $\text{Filter}(K, I)$ be the operation that filters image I with kernel K. In this operation, the filtered result has shape (H, W) and the original image is padded with 0s if necessary. Prove

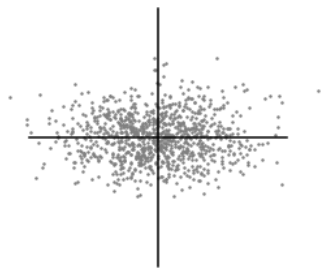
$$\text{Filter}\left(\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{bmatrix}, I\right) = \text{Filter}\left(\begin{bmatrix}1\\ 1\\ 1\end{bmatrix}, \text{Filter}(\begin{bmatrix}1 & 1 & 1\end{bmatrix}, I)\right)$$

[4 pts]

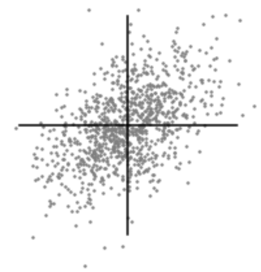
Question 6: Principal Component Analysis

(a) Draw the direction of the first principal component in figure (a). [1 pt]

(b) Draw the direction of the second principal component in figure (b). [1 pt]



(a)



(b)

(c) Suppose that we have already obtained the SVD result for some centered data matrix X (each row is the vector of a data point), which satisfies $X = U\Sigma V^\top$. By this SVD decomposition, what are the principal components and eigenvalues of PCA? [4 pts]

(d) Can you explain PCA from the optimization's perspective? What is the goal/objective of PCA [2 pts]? Is there any intuitive explanation? [2 pts]

Question 7: Harris Corner

(a) Here is the second moment matrix for Harris corner detection.[6 pts]

$$\sum_{x,y} w(x,y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}$$

When evaluated at 3 patches A, B, C of an image, the eigenvalues of the matrices are

$$A : \lambda_1 = 0.9, \lambda_2 = 0.01; B : \lambda_1 = 0.012, \lambda_2 = 0.014, C : \lambda_1 = 0.9, \lambda_2 = 0.92$$

If we know that among A, B, C, there is a corner, an edge, and a flat region. What are A, B, C respectively? Fill in the blanks below with A, B, C.

Flat region: ________. Corner: ________. Edge: ________.

(b) Fill in the the blank space below in the pseudo-code for Harris corner response. Given the second moment matrix M, the Harris corner response is computed by

$$\det(M) - 0.06(\mathrm{tr}(M))^2$$

where

$$\det\Big(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\Big) = ad - bc, \mathrm{tr}\Big(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\Big) = a + d$$

You can assume addition, subtraction, multiplication, division, and square operations are performed element-wise on multi-dimensional arrays. [4 pts]

---

function HarrisCorner(I)

$I_x \leftarrow x\_gradient(I)$

$I_y \leftarrow y\_gradient(I)$

$$W \leftarrow \begin{bmatrix} \text{---} & \text{---} & \text{---} \\ \text{---} & \text{---} & \text{---} \\ \text{---} & \text{---} & \text{---} \end{bmatrix}$$

▷ W is a 3x3 box filter

$A \leftarrow \text{convolve}(W, I_x I_x)$

$B \leftarrow \text{convolve}(W, \_\_\_\_\_\_)$

$C \leftarrow \text{convolve}(W, \_\_\_\_\_\_)$

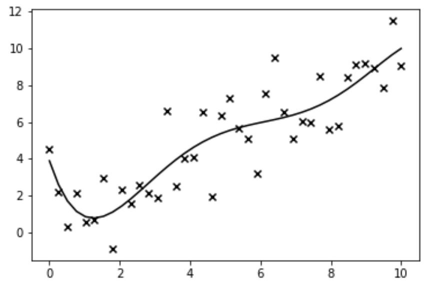
$response \leftarrow$ ________________

return $response$

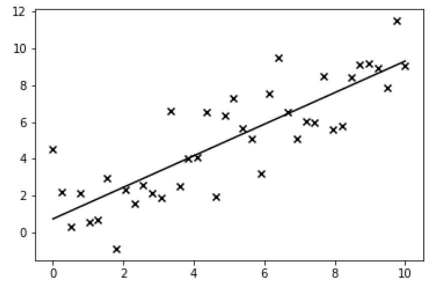
end function

Question 8: Bias and Variance

(a) The following images show two type of models for a regression task. [5 pts]



model A (K-Nearest Neighbor)



model B (Linear Regression)

Fill in the blanks (circle the correct answer) below.

Compared to Model B, Model A has ________ (lower/higher) bias and ________ (lower/higher) variance. Model A has ________ (better/worse) generalizability and is prone to ________ (overfitting/underfitting). Model A is ________ (more/less) sensitive to changes in the training data.

(b) Explain what "hyperparameter" for a model is [1 pt].

(c) Give an example of "hyperparameter" (e.g., a hyperparameter of Harris corner, KNN, K-Means, or Neural network) [1 pt].

(d) Describe the procedure to choose hyperparameters with cross-validation. (You need to list the steps in cross-validation.)[3 pts]

Question 9: Logistic Regression

Class	$S = f(x_i; w)$	Normalized Probability	$P(Y = k \mid X = x_i)$
Cat	3.4		
Dog	5.9		
Plane	-.5		

a) Explain the core function of the Softmax classifier in one to two sentences. [4 pts]

b) Given the model scores in the first column, what steps must you take in order to calculate class probability in the third column? (Explain process, do not do perform calculations) [4 pts]

c) Which class has the highest probability? [2 pts]