HOMEWORK 1:
BACKGROUND TEST

>>Jurijs Nazarovs<<
>>9075604125<<

Minimum Background Test [80 pts]

1 Vectors and Matrices [20 pts]

Consider the matrix X and the vectors $\mathbf{y}$ and $\mathbf{z}$ below:

$$X = \begin{pmatrix} 9 & 8 \\ 7 & 6 \end{pmatrix} \qquad \mathbf{y} = \begin{pmatrix} 9 \\ 8 \end{pmatrix} \qquad \mathbf{z} = \begin{pmatrix} 7 \\ 6 \end{pmatrix}$$

1. What is the inner product of the vectors $\mathbf{y}$ and $\mathbf{z}$? (this is also sometimes called the *dot product*, and is sometimes written as $\mathbf{y}^T\mathbf{z}$)
   111

2. What is the product $X\mathbf{y}$?
   $(145, 111)^T$

2 Calculus [20 pts]

1. If $y = 4x^3 - x^2 + 7$ then what is the derivative of y with respect to x?
   $12x^2 - 2x$

2. If $y = \tan(z)x^{6z} - \ln(\frac{7x+z}{x^4})$, what is the partial derivative of y with respect to x?
   $tan(z)6zx^{6z-1} - \frac{7}{7x+z} + \frac{4}{x}$

3 Probability and Statistics [20 pts]

Consider a sample of data $S = \{0, 1, 1, 0, 0, 1, 1\}$ created by flipping a coin x seven times, where 0 denotes that the coin turned up heads and 1 denotes that it turned up tails.

1. What is the sample mean for this data?
   4/7

2. What is the sample variance for this data?
   0.2857143.

3. What is the probability of observing this data, assuming it was generated by flipping a biased coin with $p(x = 1) = 0.7, p(x = 0) = 0.3$.
   $0.7^4 * 0.3^3 = 0.0064827$

4. Note that the probability of this data sample would be greater if the value of $p(x = 1)$ was not 0.7, but instead some other value. What is the value that maximizes the probability of the sample S? Please justify your answer.
   $argmax_p(p^4(1-p)^3) = 4/7$

5. Consider the following joint probability table where both A and B are binary random variables:

A	B	$P(A,B)$
0	0	0.1
0	1	0.4
1	0	0.2
1	1	0.3

(a) What is $P(A=0, B=0)$?
0.1

(b) What is $P(A=1)$?
0.2 + 0.3 = 0.5

(c) What is $P(A=0|B=1)$?
0.4/(0.4 + 0.3) = 4/7

(d) What is $P(A=0 \vee B=0)$?
0.5 + 0.3 - 0.1 = 0.7

4 Big-O Notation [20 pts]

For each pair (f, g) of functions below, list which of the following are true: $f(n) = O(g(n))$, $g(n) = O(f(n))$, both, or neither. Briefly justify your answers.

1. $f(n) = \frac{n}{2}$, $g(n) = \log_2(n)$.
   g(n) = O(f(n)), based of definition of O

2. $f(n) = \ln(n)$, $g(n) = \log_2(n)$.
   both

3. $f(n) = n^{100}$, $g(n) = 100^n$.
   f(n) = O(g(n))

Medium Background Test [20 pts]

5 Algorithm [5 pts]

Divide and Conquer: Assume that you are given a sorted array with n integers in the range $[-10, +10]$. Note that some integer values may appear multiple times in the array. Additionally, you are told that somewhere in the array the integer 0 appears exactly once. Provide an algorithm to locate the 0 which runs in $O(\log(n))$. Explain your algorithm in words, describe why the algorithm is correct, and justify its running time.

Compare 0 with the middle element of an array. If 0 is middle element, then we found it. If 0 is bigger than middle element, then new array to work with is right part from the middle element, else left part is a new array. Running time appears because in a worst case we divide our array by 2 k times. So, it would be like $n/2^k = 1$, which is $k = log_2 n$

6 Probability and Random Variables [5 pts]

6.1 Probability

State true or false. Here Ω denotes the sample space and A^c denotes the complement of the event A.

1. For any $A, B \subseteq \Omega$, $P(A|B)P(B) = P(B|A)P(A)$.
   true
2. For any $A, B \subseteq \Omega$, $P(A \cup B) = P(A) + P(B) - P(A|B)$.
   true
3. For any $A, B, C \subseteq \Omega$ such that $P(B \cup C) > 0$, $\frac{P(A \cup B \cup C)}{P(B \cup C)} \geq P(A|B \cup C)P(B \cup C)$.
   true
4. For any $A, B \subseteq \Omega$ such that $P(B) > 0, P(A^c) > 0$, $P(B|A^C) + P(B|A) = 1$.
   false
5. For any n events $\{A_i\}_{i=1}^n$, if $P(\bigcap_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$, then $\{A_i\}_{i=1}^n$ are mutually independent.
   false

6.2 Discrete and Continuous Distributions

Match the distribution name to its probability density / mass function. Below, $|\boldsymbol{x}| = k$.

(a) Laplace h

(b) Multinomial f

(c) Poisson l

(d) Dirichlet k

(e) Gamma j

(f) $f(\boldsymbol{x}; \boldsymbol{\Sigma}, \boldsymbol{\mu}) = \frac{1}{\sqrt{(2\pi)^k \det(\boldsymbol{\Sigma})}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\right)$

(g) $f(x; n, \alpha) = \binom{n}{x}\alpha^x(1-\alpha)^{n-x}$ for $x \in \{0, \ldots, n\}$; 0 otherwise

(h) $f(x; b, \mu) = \frac{1}{2b}\exp\left(-\frac{|x-\mu|}{b}\right)$

(i) $f(\boldsymbol{x}; n, \boldsymbol{\alpha}) = \frac{n!}{\Pi_{i=1}^k x_i!}\Pi_{i=1}^k \alpha_i^{x_i}$ for $x_i \in \{0, \ldots, n\}$ and $\sum_{i=1}^k x_i = n$; 0 otherwise

(j) $f(x; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}$ for $x \in (0, +\infty)$; 0 otherwise

(k) $f(\boldsymbol{x}; \boldsymbol{\alpha}) = \frac{\Gamma(\sum_{i=1}^k \alpha_i)}{\prod_{i=1}^k \Gamma(\alpha_i)}\prod_{i=1}^k x_i^{\alpha_i - 1}$ for $x_i \in (0, 1)$ and $\sum_{i=1}^k x_i = 1$; 0 otherwise

(l) $f(x; \lambda) = \lambda^x \frac{e^{-\lambda}}{x!}$ for all $x \in Z^+$; 0 otherwise

6.3 Mean and Variance

Consider a random variable which follows a Binomial distribution: $X \sim \text{Binomial}(n, p)$.

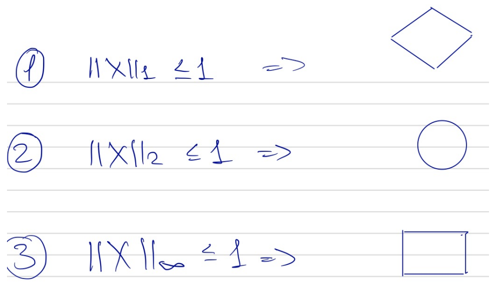
1. What is the mean of the random variable?
   np
2. What is the variance of the random variable?
   np(1-p)

7 Linear algebra [5 pts]

7.1 Norm-enclature

Draw the regions corresponding to vectors $\mathbf{x} \in \mathbb{R}^2$ with the following norms:

1. $||\mathbf{x}||_1 \leq 1$ (Recall that $||\mathbf{x}||_1 = \sum_i |x_i|$)
2. $||\mathbf{x}||_2 \leq 1$ (Recall that $||\mathbf{x}||_2 = \sqrt{\sum_i x_i^2}$)
3. $||\mathbf{x}||_\infty \leq 1$ (Recall that $||\mathbf{x}||_\infty = \max_i |x_i|$)
   Solution figure goes here.



7.2 Geometry

Prove that these are true or false. Provide all steps.

1. The smallest Euclidean distance from the origin to some point $\mathbf{x}$ in the hyperplane $\mathbf{w}^T\mathbf{x} + b = 0$ is $\frac{|b|}{||\mathbf{w}||_2}$.
   true
2. The Euclidean distance between two parallel hyperplane $\mathbf{w}^T\mathbf{x} + b_1 = 0$ and $\mathbf{w}^T\mathbf{x} + b_2 = 0$ is $\frac{|b_1 - b_2|}{||\mathbf{w}||_2}$ (Hint: you can use the result from the last question to help you prove this one).

   true

8 Programming Skills - Test Yourself

You do not need to turn this in; it is on the honor system. Please attempt the following two problems on the HankerRank website.

Grading Students (Easy): https://www.hackerrank.com/challenges/grading/problem

Verify BST (Medium): https://www.hackerrank.com/challenges/ctci-is-binary-search-tree/problem