TAKE-HOME CLASS QUIZ: DUE MONDAY OCTOBER 7: LINEAR FUNCTIONS AND EQUATION-SOLVING (PART 2)

MATH 196, SECTION 57 (VIPUL NAIK)

Your name (print clearly in capital letters): ______________________________________

PLEASE DO *NOT* DISCUSS ANY QUESTIONS EXCEPT THE STARRED OR DOUBLE-STARRED QUESTIONS.

This quiz covers some basics involving linear functions and equation-solving (notes at `Linear functions: a primer` and `Equation-solving with a special focus on the linear case`). The quiz tests for the following:

- The distinction between behavior relative to the variables (the inputs) and behavior relative to the parameters.
- Counting the number of parameters by creating the explicit general functional form from a verbal description (with a special focus on polynomial functional forms).
- Figuring out how to "ask the right questions" with respect to input choices, so that the answers provide meaningful information. This builds towards the ideas of hypothesis testing, rank, and overdetermination that we will see in the future.

(1) *Do not discuss this!*: Suppose f is a polynomial function of x of degree at most a *known number* n. What is the minimum number of (input,output) pairs that we need in order to determine f uniquely? *Extra information: Somewhat surprisingly, in this case, we do not need to be judicious about our input choices. Any set of distinct inputs of the required number will do. This has something to do with the "Vandermonde matrix" and "Vandermonde determinant" and is also related to the Lagrange interpolation formula.*
(A) $n-1$
(B) n
(C) $n+1$
(D) $2n$
(E) n^2

Your answer: ____________________________

(2) *Do not discuss this!*: f is a polynomial function of two variables x and y of total degree at most 2. In other words, for each monomial occurring in f, the total of the degrees of x and y in that monomial is at most 2. No other information is given about f. What is the minimum number of judiciously chosen (input,output) pairs we need in order to determine f uniquely?
(A) 2
(B) 3
(C) 4
(D) 6
(E) 7

Your answer: ____________________________

(3) *Do not discuss this!*: f is a polynomial function of two variables x and y of total degree at most 3. In other words, for each monomial occurring in f, the total of the degrees of x and y in that monomial is at most 3. No other information is given about f. What is the minimum number of judiciously chosen (input,output) pairs we need in order to determine f uniquely?
(A) 3
(B) 6
(C) 8
(D) 9
(E) 10

Your answer: ____________________________

(4) (*) *The perils of overfitting; see also Occam's Razor*: Suppose we are trying to model a function that we expect to behave in a polynomial-like manner, though we don't really have a good reason to believe this. Additionally, there is a possibility for measurement error in our observations. Our goal is to find the parameters so that we can both predict unmeasured values and do a qualitative analysis of the nature of the function and its derivatives and integrals.

We have a large number of observations (say, several thousands). We could attempt to "fit" the function using a polynomial of degree n for some fixed n using all those data points, and we will get a certain "best fit" that minimizes the deviation between the curve used for fitting and the function being fit. For instance, for $n = 1$, we are trying to find the best fit by a straight line function. For $n = 2$, we are trying to find the best fit by a polynomial of degree at most 2. We could try fitting using different values of n. Which of the following is true?

If you are interested in more on this, look up "overfitting". A revealing quote is by mathematician and computer scientist John von Neumann: "With four parameters I can fit an elephant. And with five I can make him wiggle his trunk." Another is by prediction guru Nate Silver: "The wide array of statistical methods available to researchers enables them to be no less fanciful and no more scientific than a child finding animal patterns in clouds."
(A) Larger values of n give better fits, therefore the larger the value of n we use, the better.
(B) Smaller values of n give better fits, therefore the smaller the value of n we use, the better.
(C) Larger values of n give better fits, therefore the larger the value of n we use, the less impressive a good fit (i.e., low deviation between the polynomial and the actual set of observations) should be.
(D) Smaller values of n give better fits, therefore the smaller the value of n we use, the less impressive a good fit (i.e., low deviation between the polynomial and the actual set of observations) should be.
(E) The value of n we use for trying to get a good fit is irrelevant. A good fit is a good fit, regardless of the type of function used.

Your answer: ____________________________

(5) (*) F is an affine linear function of two variables x and y, i.e., it has the form $F(x, y) := ax + by + c$ with a, b, and c real numbers. We want to determine the values of the parameters a, b, and c by using input-output pairs. It is, however, costly to find input-output pairs. We have already found $F(1, 3)$ and $F(3, 7)$. We want to find F for one other pair of inputs to determine a, b, and c. Which of these will *not* be a good choice?
(A) $F(2, 2)$, i.e., the input $x = 2$, $y = 2$
(B) $F(2, 3)$, i.e., the input $x = 2$, $y = 3$

(C) $F(2,4)$, i.e., the input $x = 2$, $y = 4$
(D) $F(2,5)$, i.e., the input $x = 2$, $y = 5$
(E) $F(2,6)$, i.e., the input $x = 2$, $y = 6$

Your answer: ____________________________