Data 100, Spring 2023

Discussion #1

Welcome to Data 100!

Linear Algebra Fundamentals

1. Linear algebra is what powers linear regression, logistic regression, and PCA (concepts we will be studying in this course). Moving forward, you will need to understand how matrix-vector operations work. That is the aim of this problem.

   Alice, Bob, and Candace are shopping for fruit at Berkeley Bowl. Berkeley Bowl, true to its name, only sells fruit bowls. A fruit bowl contains some fruit and the price of a fruit bowl is the total price of all of its individual fruit.

   Berkeley Bowl has apples for \$2, bananas for \$1, and cantaloupes for \$4. (expensive!). The price of each of these can be written in a vector:

$$\vec{v} = \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix}$$

   Berkeley Bowl sells the following fruit bowls:

   1. 2 of each fruit
   2. 5 apples and 8 bananas
   3. 2 bananas and 3 cantaloupes
   4. 10 cantaloupes

   (a) Define a matrix B such that $B\vec{v}$ evaluates to a length 4 column vector containing the price of each fruit bowl. The first entry of the result should be the cost of fruit bowl 1, the second entry the cost of fruit bowl 2, etc.

(b) Alice, Bob, and Candace make the following purchases:

- Alice buys 2 fruit bowl 1s and 1 fruit bowl 2.
- Bob buys 1 of each fruit bowl.
- Candace buys 10 fruit bowl 4s (he really like cantaloupes).

Define a matrix A such that the matrix expression $AB\vec{v}$ evaluates to a length 3 column vector containing how much each of them spent. The first entry of the result should be the total amount spent by Alice, the second entry the amount sent by Bob, etc.

(c) Let's suppose Berkeley Bowl changes their fruit prices, but you don't know what they changed their prices to. Alice, Bob, and Candace buy the same quantity of fruit baskets and the number of fruit in each basket is the same, but now they each spent these amounts:

$$\vec{x} = \begin{bmatrix} 80 \\ 80 \\ 100 \end{bmatrix}$$

In terms of A, B, and $\vec{x}$, determine $\vec{v_2}$ (the new prices of each fruit).

(d) In the previous part, we assumed that AB (the matrix multiplication of A and B) was invertible. Why is AB (as calculated below) invertible? State two conditions for an arbitrary matrix to be invertible.

$$AB = \begin{bmatrix} 9 & 12 & 4 \\ 7 & 12 & 15 \\ 0 & 0 & 100 \end{bmatrix}$$

2. As a warm up for the homework, we will introduce matrix inverses and matrix rank.

- The inverse of a square invertible matrix M, M^{-1} is defined as a matrix such that $MM^{-1} = I$ and $M^{-1}M = I$. The matrix I is a special matrix denoted as the identity matrix where the diagonal elements are 1 and the non-diagonal elements are 0.

- Linear dependence among a set of vectors $\{v_1, v_2, v_3, ..., v_n\}$ is defined as follows. If any (non-trivial) linear combination of the vectors can produce the zero vector, then the set of vectors is linearly dependent.

  In other words, if we can multiply the vectors v_i with some scalar α_i and sum the quantity to obtain the zero vector (given at least one $\alpha_j \neq 0$, then the set is linearly dependent.

$$\sum_{i=1}^{n} \alpha_i v_i = 0 \text{ such that some } \alpha_j \neq 0 \implies \text{linear dependence}$$

  Any set of vectors such that we cannot obtain the zero vector as described above is linearly independent.

- The (column) rank of a matrix M is the maximal number of linearly independent column vectors in M. A full rank matrix has a column rank equal to the number of column vectors.

We will go over all of these definitions applied to relevant practical examples in the following subparts.

(a) Consider the matrix $M = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} v_1 & v_2 \end{bmatrix}$ containing two column vectors $v_1 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$ and $v_2 = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$. Is it possible to construct the zero vector using a linear combination of the column vectors? What can be concluded about the rank of the matrix M?

(b) Consider the inverse matrix $M^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ of M. Carry out the matrix multiplication MM^{-1}, and determine what M^{-1} must be.

(c) Consider a different matrix $Q = \begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & 5 \end{bmatrix} = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix}$. What is the column rank of the matrix? Is the matrix invertible?

(d) Consider a matrix R, which is equal to the transpose of the matrix Q: $R = Q^T$. What is the column rank of the matrix R? Is the matrix R invertible?

Calculus

In this class, we will have to determine which inputs to a functions minimize the output (for instance, when we choose a model and need to fit it to our data). This process involves taking derivatives.

In cases where we have multiple inputs, the derivative of our function with respect to one of our inputs is called a *partial derivative*. For example, given a function $f(x, y)$, the partial derivative with respect to x (denoted by $\frac{\partial f}{\partial x}$) is the derivative of f with respect to x, taken while treating all other variables as if they're constants.

3. Suppose we have the following scalar-valued function on x and y:

$$f(x, y) = x^2 + 4xy + 2y^3 + e^{-3y} + \ln(2y)$$

(a) Compute the partial derivative of $f(x, y)$ with respect to x.

(b) Compute the partial derivative of $f(x, y)$ with respect to y.

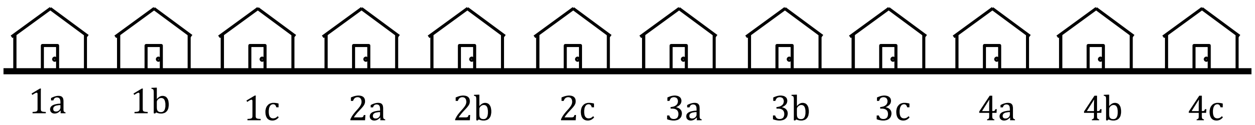
(c) The gradient of a function $f(x, y)$ is a vector of its partial derivatives. That is,

$$\nabla f(x, y) = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix}^T$$

$\nabla f(x, y)$ tells us the magnitude and direction in which f is moving, at point (x, y). This is analogous to the single variable case, where $f'(x)$ is the rate of change of f, at the point x.

Using your answers to the above two parts, compute $\nabla f(x, y)$ and evaluate the gradient at the point $(x = 2, y = -1)$.

Probability & Sampling



4. Kalie wants to measure interest for a party on her street. She assigns numbers and letters to each house on her street as illustrated above. She picks a letter "a", "b", or "c" at random and then surveys every household on the street ending in that letter.

   (a) What is the chance that two houses next door to each other are both in the sample?

   (b) Now, suppose that Kalie decides to collect an SRS of one house instead. What is the probability that house 1a is **not** selected in Kalie's SRS of one house?

   (c) Kalie decides to collect a SRS of four houses instead of a SRS of one house. What is the probability that house 1a is **not** in Kalie's simple random sample of four houses?

   11/12 * 10/11 * 9/10 * 8/9 = 8/12 = 2/3

   (d) Instead of surveying every member of each house from the SRS of four houses, Kalie decides to only survey two members in each house. Four people live in house 1a, one of whom is Bob. What is the probability that Bob is **not** chosen in Kalie's new sample?

   1/3 * 1/2 = 1/6; 1-1/6 = 5/6