Assignment 0

Foundations of Machine Learning
IIT-Hyderabad
Aug-Dec 2021

This homework is intended to review basic pre-requisites in the following topics:

- Linear Algebra, Probability, Python programming

1 Practice Questions

The questions below are only for your practice - no submission required. (The exam could include questions from this list though!)

1. **Numpy:** The following questions should be done using `numpy` operations. Each exercise can be solved with a maximum of 4-5 lines of Python code. Please avoid the use of iterative constructs (such as for loops) to the extent possible, and use matrix/vector operations to achieve the objectives.

   (a) Import the numpy package under the name np. Print the numpy version and the configuration.

   (b) Create a null vector of size 10, and output the vector to the terminal.

   (c) Create a null vector of size 10 but the fifth value which is 1. Output the vector to the terminal.

   (d) Reverse a vector (first element becomes last).

   (e) Create an $n \times n$ array with checkerboard pattern of zeros and ones.

   (f) Given an $n \times n$ array, sort the rows of array according to m^{th} column of array.

   (g) Create an $n \times n$ array with $(i+j)^{th}$-entry equal to $i+j$.

   (h) Consider a $(6, 7, 8)$ shape array, what is the index (x, y, z) of the 100th element (of the entire structure)?

   (i) Multiply a 5×3 matrix by a 3×2 matrix (real matrix product).

   (j) Create random vector of size 10 and replace the maximum value by 0.

   (k) How to find the closest value (to a given scalar) in an array?

   (l) Subtract the mean of each row from each corresponding row of a matrix.

   (m) Consider a given vector, how to add 1 to each element indexed by a second vector (be careful with repeated indices - you should consider it only once)?

   (n) How to find the most frequent value in an array?

(o) Extract all the contiguous 3×3 blocks from a random 10×10 matrix.

(p) Compute the rank, trace and determinant of a matrix.

2. **Linear Algebra:**

- Prove or disprove: Empty set is a vector space.
- - Show that the inverse of $M = I + (\mathbf{u}\mathbf{v}^T)$ is of the type $I + \alpha(\mathbf{u}\mathbf{v}^T)$, where $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, $\mathbf{v}^T\mathbf{u} \neq 0$.
  - Continuing from previous question, find α.
  - For what $\mathbf{u}$ and $\mathbf{v}$ is M singular?
  - Find the null space of M, if it is singular.
- - Consider the 2×2 matrix:
  $$A = \begin{bmatrix} -2 & 2 \\ -6 & 5 \end{bmatrix}$$
  There exists vectors such that when the matrix A acts on those vectors, the subspace of the vectors does not change. Mathematically, $Ax = \lambda x$. The vectors x are called *eigenvectors* and the values λ are the corresponding *eigenvalues*. Find the eigenvalues and corresponding eigenvectors for the matrix A.
  - Consider a diagonal matrix Λ which has the eigenvalues of A as its diagonal entries. Find the matrix U such that the equation $AU = U\Lambda$ holds.
  - Note that we can write the matrix A as $A = U\Lambda U^{-1}$. Find the inverse of the matrix U computed in the previous question and verify.
- Show that for any square matrix A, the eigenvectors of A are also eigenvectors of A^2. What are the eigenvalues for A^2?

3. **Probability:**

- If two binary random variables X and Y are independent, are $\bar{X}$ ($\bar{X}$ is the complement of X) and Y also independent? Prove your claim.
- Show that if two variables x and y are independent, then their covariance is zero.
- By using a change of variables, verify that the univariate Gaussian distribution given by:
$$\mathcal{N}(x|\mu,\sigma^2) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left(- \frac{1}{2\sigma^2}(x-\mu)^2 \right)$$
satisfies the equation:
$$\mathbb{E}[x] = \int_{-\infty}^{\infty} \mathcal{N}(x|\mu,\sigma^2) x dx = \mu$$
Next, by differentiating both sides of the normalization condition:
$$\int_{-\infty}^{\infty} \mathcal{N}(x|\mu,\sigma^2) dx = 1$$
with respect to σ^2, verify that the Gaussian satisfies the equation:
$$\mathbb{E}[x^2] = \int_{-\infty}^{\infty} \mathcal{N}(x|\mu,\sigma^2) x^2 dx = \mu^2 + \sigma^2$$

- There are two coins $C1$ and $C2$. $C1$ has a equal prior on a head ($H = 1$) or tail ($T = 0$) and the fate of $C2$ is dependent on $C1$. If $C1$ is a head, $C2$ will be a head with probability 0.7. If $C1$ is a tail, $C2$ will be a head with probability 0.5. $C1$ and $C2$ are tossed in sequence once, and the observed sum of the two coins, $S = C1 + C2$, is 1. What is the probability that $C1 = T$ and $C2 = H$ (*Hint: use Bayes theorem*)?
- Suppose that we have three coloured boxes r (red), b (blue), and g (green). Box r contains 3 apples, 4 oranges, and 3 limes, box b contains 1 apple, 1 orange, and 0 limes, and box g contains 3 apples, 3 oranges, and 4 limes. If a box is chosen at random with probabilities p(r) = 0.2, p(b) = 0.2, p(g) = 0.6, and a piece of fruit is removed from the box (with equal probability of selecting any of the items in the box), then what is the probability of selecting an apple? If we observe that the selected fruit is in fact an orange, what is the probability that it came from the green box?