EE 524 Machine Learning Lab

Assignment 2

20 August 2021

1 Probability distributions

- **Continuous distributions**

  (a) Plot the PDF and CDF for the following distributions. Vary the parameters to get at least 3 different realizations of each: Uniform (a, b), Exponential (λ), Gamma (γ), Beta (β), Gaussian (μ, σ), Standard Normal $(0, 1)$, Chi-squared (k). How many of these distributions are non negative? Calculate the mean, median, mode and variance and plot it with the PDF in a single plot for one distribution. Also draw the CDFs separately.

  (b) Consider a signal with 2 orthogonal components X, Y . Each component is a random variable that is independent of the other and follows Gaussian distribution with 0 mean and equal variances $\sigma^2 = 4$. Can you find the distribution of the power content of the signal i.e. Power $= X^2 + Y^2$. Which distribution is this? Is it a non negative distribution? Compute the mean, median, mode and variance and plot with the PDF.

- **Discrete distributions**

  (a) Consider a transmitter that can transmit messages only in the form of 2 bits 0 and 1. The receiver receives a single message sent by the transmitter say X. Since, we do not know before hand which bit was sent X follows a certain distribution. Which distribution is this? Find the PMF and plot it. Also calculate the mean and variance of X. Assume probability for receiving 0(failure) is $p = 0.1, 0.5, 0.8$.

  (b) Consider the same transmitter but this time the receiver receives 20 bits. Each bit is independent of the previous one. The probability of receiving a 0 is $p = 0.4$. Let X be the number of 1s (successes) that is received at the receiver's end. What is the distribution followed by X. How is it related to the first distribution? Calculate the mean and variance and plot the PMF.

(c) Consider an experiment in which bits are transmitted sequentially. Each bit is independent of the other. When the first bit arrives the receiver checks for a 1. If it is 1 the receiver takes a decision. If it is a 0 the receiver waits for the next bit. This goes on till the receiver finds out that the bit was 1 and the experiment is dropped once the bit is 0. Let X denote that the bit received was 1 on the k^{th} time (example, $X = 2$ suggests that the bit sequence was 0, 1 and once 1 was received the experiment was stopped). Find the PMF of X. What is this distribution known as? Calculate the mean and variance and plot the PMF. How is this distribution related to the first distribution.

(d) Consider an experiment in which the receiver receives different bits randomly from many transmitters within a second's interval. Each second it was observed that each second (at any interval) the mean number of bits that were received was 5. Let X be the number of bits received every second (during an interval). What is the distribution that is followed by X. Is there any relation between this and the second distribution that we discussed? Calculate the mean, variance and plot the PMF.

2 Linear Algebra and Optimization

(a) A matrix is said to be positive semi definite (PSD) when $x^T Ax \geq 0$ for all $x \epsilon R^n$. If the eigen values are all greather than or equal to 0 then also the matrix is said to be PSD. Find the eigen values and eigen vectors of the following matrices:

$$\begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}, \begin{bmatrix} 9 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 9 \end{bmatrix}, \begin{bmatrix} 9 & 5 \\ 5 & 4 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -0.5 \\ -0.5 & 1 \end{bmatrix},$$
$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$$

Take a 2 dimensional vector x = $[x1 \;\; x2]$ and use the quadratic form $x^T Ax$ to find the equation and plot the curve. Find eigen vectors and use them as the orthogonal axes for the plot. Then use the square root of the eigen values to see till what extent the curve will go.

What do you think happens if we have a matrix of higher dimension? What will be the shape of a quadratic form from a n-dimensional PSD matrix?

(b) A co-variance matrix is said to be a PSD matrix. Check whether the following are valid co-variance matrices? Also plot the accompanying quadratic form for the matrices:

$$\begin{bmatrix} -1 & 2 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix}, \begin{bmatrix} 2 & 5 \\ 5 & 7 \end{bmatrix}$$

(c) In question (a), calculate the gradients (https://en.wikipedia.org/wiki/Gradient) of the quadratic forms and compute the stationary points (by hand, no code required).

(d) Hessian matrix is a square matrix of second order derivatives of a scalar-valued function. More on this here https://en.wikipedia.org/wiki/Hessian_matrix. There is a relation between Hessian, positive semi-definiteness and convexity of a curve. Compute the quadratic forms of the matrices given below and then compute the Hessian Matrix:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 9 \end{bmatrix}, \begin{bmatrix} 2 & 4 \\ 4 & 6 \end{bmatrix}, \begin{bmatrix} 2 & 6 \\ 6 & 7 \end{bmatrix}, \begin{bmatrix} 2 & 6 \\ 6 & 18 \end{bmatrix}, \begin{bmatrix} 2 & 6 \\ 6 & 20 \end{bmatrix}$$

Also plot the quadratic form in the 3D space and try to draw conclusions between the minima/maxima and the eigen values of Hessian.

(e) A Hessian matrix of a 2-dimensional feature space is given as $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$. Comment about the possible quadratic form. Also comment about the convexity from the hessian matrix. Can you plot the quadratic form?

(f) Read the iris file (provided with the assignment) into a matrix with all the features (attributes). Each column of matrix should represent a vector. Compute the co-variance matrix of the vectors and find the eigen values and eigen vectors, print them in descending order.