CS771A/CS771 - Assignment #0
Pre-requisites: Probability and linear algebra

January 6, 2016

1. *This is an assignment to revise your pre-requisites.*
2. *It will not be graded but you have to submit it. Other assignments will be given in parallel which will be graded.*
3. *It is due by 30th Jan. 16, 17.00. Upload it to the asn1 directory on the ftp site. Your file name should by your rollno followed by Asn0. Example: if your roll number is 12345 then your file name will be 12345Asn0.pdf.*
4. *Submit only PDF files. Latex preferred but scanned handwritten files also allowed.*

1. Prove the following inequalities, where $C_1, C_2, ..., C_n$ denote subsets (events) from the sample space C.

   (a)
   $$\sum_{i=1}^{n} P(C_i) - \sum_{1 \leq i < j \leq n} P(C_i \cap C_j) \leq P(\bigcup_{i=1}^{n} C_i) \leq \sum_{i=1}^{n} P(C_i)$$

   (b)
   $$P(\bigcap_{i=1}^{n} C_i) \geq \sum_{i=1}^{n} P(C_i) - n + 1$$

   (c)
   $$P(\bigcap_{i=1}^{n} C_i^c) \leq exp(-\sum_{i=1}^{n} P(C_i))$$

   For each case, give an example where the equality holds.

2. (a) Suppose the joint probability density function (p.d.f.) of (X,Y) is given by :
   $$f(\text{x,y}) = \begin{Bmatrix} cx^2y \text{ when } 0 < x < y < 1, \\ 0 \text{ otherwise.} \end{Bmatrix}$$

   (a) Find the value of c

   (b) Find the marginal pdf's of X and Y

   (c) Determine whether X is independent to Y

   (d) Calculate $P((X + Y) < 1)$

   (b) Let the p.d.f of X be given as:
   $$f(x) = \begin{Bmatrix} \sin(\text{x}) \text{ when } 0 < x < \frac{\pi}{2}, \\ 0 \text{ otherwise.} \end{Bmatrix}$$

(a) Find the moment generating function $M(t)$ of $f(x)$.

(b) Compute $E[x^3]$

(c) Let X and Y be random variables such that E[X] = 4, E[Y] = -4, E[X^2] = E[Y^2] = 20, correlation(X,Y) = -0.5. Find the values of the following expressions:

(a) E[2X - Y]

(b) var(X + Y)

(c) E($(X - Y)^2$)

3. Let $X_1, X_2, X_3, ..., X_n$ be independent random variables, each with a Uniform distribution over (0, 1). Find the value of following :

(a) E[max$(X_1, X_2, ..., X_n)$]

(b) E[min$(X_1, X_2, ..., X_n)$]

(c) var(max$(X_1, X_2, ..., X_n)$)

(d) var(min$(X_1, X_2, ..., X_n)$)

Hint : use cumulative distribution function to first find the pdf for max$(X_1, X_2, ..., X_n)$ *and min*$(X_1, X_2, ..., X_n)$

4. (a) Let X be a non-negative random variable. Under the assumptions that E[X] exists, prove that for any positive constant c,

$$P[X \geq C] \leq \frac{E[X]}{c}$$

Is this a tight Inequality? Give examples / situations where the equality holds.

(b) Let X be a random variable with a finite variance σ^2. (A finite variance implies a finite mean, let us denote it by μ). Using the result in part (a), prove that :

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

(c) Let X be a random variable such that E[X] = 3 and E[X^2] = 13, Using the results in part (a) and (b), find a lower bound for P(-2 < X < 8).

5. (a) Given that

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

Compute $A^3 - 15A^2 - 17A$.

(b) Given that

$$A = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$$

Compute A^5. How will you efficiently compute A^k for any positive integer k.

6. Let A_1, A_2 and A_3 be mutually independent events. Also, $P[A_1] = \frac{1}{2}$, $P[A_2] = \frac{1}{3}$ and $P[A_3] = \frac{1}{4}$. Compute $P[(A_1 \cap A_2) \cup A_3^c]$.

7. With probability 0.9, a spam filter can detect a spam mail when applied on spam mails, and with probability 0.15 it gives false positive when applied to a non-spam mail. Estimation is that 11% of the mails are spams. Suppose that the spam filter is applied on a new mail of which we have no prior information of being a spam or non-spam. Compute the following probabilities:
   (a) that the mail is detected as a spam;
   (b) that, given that the mail is detected as a spam, the mail is a spam;
   (c) that, given that the mail is detected as non-spam, the mail is a non-spam;
   (d) that the mail is misclassified.
8. Calculate the derivative of $log|X|$ wrt X. Assume X to be a symmetric matrix with positive determinant value (denoted by $|.|$). Using the result, derive the expression for the gradient for the following function $f : \mathcal{R}^n \rightarrow \mathcal{R}$ such that,
$$f(\mathbf{x}) = log|F_0 + x_1 F_1 + \cdots + x_n F_n|$$
where F_i are $p \times p$ symmetric matrices. You can assume that the domain of $\mathbf{x}$ and form of F_i are such that the log function is defined.
9. (a) Let $\mathbf{X}_{3\times 1}$ be $N_3(\boldsymbol{\mu}, \Sigma)$ with
$$\Sigma = \begin{bmatrix} 4 & 1 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$
Are X_1 and X_2 independent? What about (X_1, X_2) and X_3?
   (b) Let $X = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix}$ be distributed as $N_p(\boldsymbol{\mu}, \Sigma)$ with $\begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}$, $\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$. Show that the conditional distribution of $\mathbf{X}_1$, given that $\mathbf{X}_2 = \mathbf{x}_2$ is normal and has
$$Mean = \boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)$$
and
$$Covariance = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$$
   (c) The joint probability density of a normally distributed $n \times 1$ random vector $\mathbf{Y}$ is given by
$$f(\mathbf{y}) = \frac{1}{(2\pi)^{n/2}}|\Sigma|^{-1/2} exp\{-(\mathbf{y} - \boldsymbol{\mu})'\Sigma^{-1}(\mathbf{y} - \boldsymbol{\mu})/2\}$$
where $\boldsymbol{\mu}$ is the mean vector and Σ is the covariance matrix. Considering a bivariate normal distribution, the correlation between Y_1 and Y_2 is given by $\rho_{12} = \frac{\sigma_{12}}{\sigma_{11}\sigma_{22}}$. Evaluate the form of probability distribution for this case. Further, describe the form of density plot on the 2-d plane for the cases when Y_1 and Y_2 are positively correlated, negatively correlated and not associated. Note that each case mentioned above corresponds to the sign of the correlation value. Use figures to indicate the pattern of density plot for each case.
10. (a) Show that for a real symmetric matrix, the eigenvalues form an orthonormal basis.
   (b) Given a set of random variables $(x_1, x_2, \ldots, x_n)$, the covariance matrix is defined so that $A_{i,j} = Cov(x_i, x_j)$. Show that for two vectors $\mathbf{a}$ and $\mathbf{b}$, the values $\langle \mathbf{a}, \mathbf{x} \rangle$ and $\langle \mathbf{b}, \mathbf{x} \rangle$, are correlated as $Cov(x_a, x_b) = \mathbf{a}' A \mathbf{b}$.

(c) Using the above definition, one can define the variance in a particular direction for the values. Find the direction in the $\mathcal{R}^n$ space which has the greatest variance. [Note: We would take up a very important application of this result further in the course.]