Middle East Technical University
Department of Computer Engineering

CENG 564

PATTERN RECOGNITION

Fall 2018

TAKE-HOME EXAM 1

The aim of this assignment is to introduce some of the basic mathematical tools you will be using in Pattern Recognition. The mathematical prerequisites are a solid background in calculus, probability and linear algebra. You are expected to give rigorous solutions, clearly identifying and justifying every step. You can refer to the appendices at the back of the textbook or other widely available sources in case you cannot remember the exact definitions, theorems etc. You can make use of any source for solving a question; however your exposition should demonstrate your understanding in a rigorous manner and you should properly cite your sources.

QUESTIONS

PART 1 : Probability Theory

	Y = 10		Y = 20		Y = 30	
	Z = 0	Z = 1	Z = 0	Z = 1	Z = 0	Z = 1
X = 3	0.025	0.025	0.03	0.02	0.05	0.15
X = 5	0.075	0.050	0.025	0.030	0.020	0.2
X = 7	0.04	0.06	0.025	0.050	0.025	0.1

1) Answer the questions below according to the **joint probability** table above and you have to show your work by the formulas of **CONDITIONAL PROBABILITY, MARGINALIZATION**, otherwise answers will be severely penalized.

i. P(Z = 1 | X = 7)
ii. Show the marginal probability distribution of P(X).
iii. Using the distribution found in **b,** find $E_X[x]$ and $Var_X[x]$
iv. Show the marginal probability distribution of P(X | Z = 1).
v. Using the distribution found in **d,** find $E_{X|Z=1}[x]$ and $Var_{X|Z=1}[x]$

2) Suppose you choose a *real number* X randomly from the interval $[10, 100]$.

i. Find the density function $f(x)$ and the probability of an event I for this experiment, where I is a subinterval $[a, b]$ of $[10, 100]$.
ii. Find the expected value $E_X[x]$
iii. From (i) find the probability of $X^2 - 110X + 2800 > 0$

iv. What if we have chosen the real number with the density function, $f(x) = Cx$, rather than uniform, find C, and expected value $E_X[x]$.

3) Let $X \sim U(-1,1)$, find pdfs for the following:

i. $\sqrt{(|X|)}$
ii. $-ln|X|$

4) Let XY be two uniformly distributed random variables indicating a coordinate on a triangle $T:((0,1),(1,0),(0,-1))$, find both cumulative and probability density function of $|X-Y|$.

5) Prove the following equations and assume that X and Y are random variables (RVs) with joint probability distribution $P(X,Y)$:

i. $E[X^2] - E[X]^2 = E[(X - E[X])^2]$; X and Y are continuous RVs
ii. $E[X] = E_Y[\, E_{X|Y}[\, X \mid Y]\,]$; X and Y are discrete RVs
iii. $X - Y$ and $X + Y$ are uncorrelated ; X and Y are discrete with same variance

6) Show that;

i. independence implies uncorrelation.
ii. uncorrelation does not imply independence.
iii. Pairwise independence does not (always) imply independence

7) Suppose we randomly and independently pinpointed two points on the unit interval with uniform distribution. Find the expected distance between the selected points.

8) Suppose we have a bag of balls numbered from 1 to 20, and we are going to do a selection with replacement until 20 comes up.

i. What is the probability that we perform the selection for an arbitrary number, n, of times ?
ii. Find the expected number of selections.

9) Suppose you have a bivariate Gaussian distribution, $(x_1, x_2) \sim N(\begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix}, \begin{bmatrix}\sigma_1^2 & \sigma_1\sigma_2\\ \sigma_2\sigma_1 & \sigma_2^2\end{bmatrix})$. For given different values of mean and covarince matrices, inspect and sketch 2D projections of Gaussian distributions. You can use MATLAB. (*It is advised to inspect more than the given on your own, until you are comfortable with Gaussian distribution*) *i.* $\begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix} = \begin{bmatrix}1\\ 2\end{bmatrix}, \begin{bmatrix}\sigma_1^2 & 0\\ 0 & \sigma_2^2\end{bmatrix}, \qquad \sigma_1^2 > \sigma_2^2$

ii. $\begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix} = \begin{bmatrix}1\\ 2\end{bmatrix}, \begin{bmatrix}\sigma_1^2 & 0.5\\ 0.5 & \sigma_2^2\end{bmatrix}, \qquad \sigma_1^2 = \sigma_2^2$

PART 2: Linear Algebra & Matrix Calculus

10) Suppose that $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n$ form a basis of R^n and $\boldsymbol{y} = \lambda_1\boldsymbol{a}_1 + \lambda_2\boldsymbol{a}_2 + \dots + \lambda_n\boldsymbol{a}_n$ with $\lambda_j = 0$. Prove that $\boldsymbol{a}_1, \dots, \boldsymbol{a}_{j-1}, \boldsymbol{y}, \boldsymbol{a}_{j+1} \dots, \boldsymbol{a}_n$ do not form a basis of R^n.

11) Prove that if $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$ spans V, then so does the list

$$\{\boldsymbol{v}_1 - \boldsymbol{v}_2, \boldsymbol{v}_2 - \boldsymbol{v}_3, \dots, \boldsymbol{v}_{n-1} - \boldsymbol{v}_n, \boldsymbol{v}_n\}$$

obtained by subtracting from each vector the succeding vector, except the last one.

12) Given the matrix A below ;

$$A = \begin{bmatrix} 5 & 4 & 2\\ 4 & 5 & 2\\ 2 & 2 & 2\end{bmatrix}$$

i. Find the eigenvalues & eigenvectors by finding the characteristic equation.

ii. Generate the diagonalization formula of A step by step by changing the coordinate system using an ***orthogonal*** basis, B, being composed of ***orthogonal*** eigenvector set.

13) A positive definite matrix $\boldsymbol{A}$ can be defined as the matrix making the following always positive for any real value of the vector $\boldsymbol{x}$;

$$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}$$

Show that a necessary and non-sufficient condition for $\boldsymbol{A}$ to be positive definite is that all of the eigenvalues of $\boldsymbol{A}$ are positive.

14) Let *f* be a function $f : R^{mxn} \rightarrow R$, A be a matrix of size $m \ x \ n$, x is a vector $x \ \in R^n$.

Prove the following *Gradient* and *Hessian* equations :

i. $\nabla_x(x^T x) = 2x$
ii. $\nabla_x(x^T Ax) = 2Ax \ (if \ A \ is \ symmetric)$
iii. $\nabla_x^2(x^T Ax) = 2A \ (if \ A \ is \ symmetric)$

15) True/False. Prove or disprove the following;

i. For every NxM matrix, the number of linearly independent columns and rows are equivalent.
ii. Eigenvectors of a symmetric matrix with real values always form an orthonormal basis.
iii. A real matrix has orthogonal eigenvectors if and only if $A^T A = AA^T$
iv. Two distinct eigenvectors corresponding to the same eigenvalue are always linearly DEPENDENT.
v. Eigenvalues cannot be 0.
vi. If an NxN matrix, A, has linearly dependent columns, then the linear transformation, $T(x) = Ax$, is a 1-1 and onto linear mapping between the original space R^N and the new space (linearly) transformed by this matrix.

BONUS

16) We know the implication relationship between correlation and causation. Suppose that we check for the correlation between craving (C) and birth (B) and found out that $P(B|C) > P(B)$. Now, introduce another factor to the picture binary variable pregnancy (P) which we get $P(B|C, P = true) = P(B|P = true)$ and $P(B|C, P = false) = P(B|P = false)$. Based on this findings, show what *might* be the causal link in this picture by determining whether B and C are conditionally independent given P.
Now, after gaining insight from previous picture, you need to devise a different scenario which might seem as a paradox. Suppose, that there is a positive correlation between kidney disease (K) and protein consumption (C), $P(K|P) > P(K)$, which makes it seem as protein consumption causes kidney problems. Now, imagine we bring gender (G) to the picture and find out that $P(K| \ P, \ G = female) < P(K \ |G = female)$ and $P(K| \ P, \ G = male) < P(K \ |G = male)$. What would this reversal of correlations indicate? Explain this reversal by setting up an actual scenario by generating your populations using set scheme and give the appropriate proportions in detail that will lead to this "*paradox*".