Course: Machine Learning - Foundations
Week 1 (Test questions)

---

1. (1 point)

   Answer: $[2, 4, -5] \in R^3$

2. (1 point)

   Answer: C

   It may be appropriate for classification but not for regression as it uses a fixed penalty of 1 if the output does not match with the target value, which may not be a useful loss calculation.

3. (1 point)

   Answer: C

   Target variable is discrete

4. (1 point)

   Answer: C

   remainder is 1

5. (1 point)

   Answer: C

   labels are absent for unsupervised learning

6. (1 point)

   Answer: C

7. (1 point)

   Answer: B,D

8. (1 point)

   Answer: B

9. (1 point)

   Answer: C

   Training set is used to fit the model validation set is used for model selection and test set is used for evaluating the final model

10. (1 point)

    Answer: C

    $\dfrac{1}{n}\sum_{i=1}^{n} -log(P(X^i))$ is used for density estimation, $\dfrac{1}{n}\sum_{i=1}^{n} ||g(f(X^i)) - X^i||^2$ is used for dimansionality reduction, $\dfrac{1}{n}\sum_{i=1}^{n} ||f(X^i) - Y^i||^2$ is used for regression and $\dfrac{1}{n}\sum_{i=1}^{n} \mathbf{1}(f(X^i) \neq Y^i)$ is used for classification

11. (1 point)

Answer: Pair 1: 15.225 [Range could be 15 to 16]
Pair 2: 3.8 [Range could be 3 to 4] Pair 1: $Loss = \frac{0.625+10.625+19.225+30.425}{4} = 15.225$

x_1	x_2	$f(\mathbf{x})$	$g(f(x^i)) - x^i$	$\|\|g(f(x^i)) - x^i\|\|^2$
1	0.5	0.5	[-0.75, -0.25]	0.625
2	2.3	-0.3	[-2.15, -2.45]	10.625
3	3.1	-0.1	[-3.05, -3.15]	19.225
4	3.9	0.1	[-3.95, -3.85]	30.425

Pair 2: $Loss = \frac{0.406+2.356+4.656+7.80}{4} = 3.8$

x_1	x_2	$f(\mathbf{x})$	$g(f(x^i)) - x^i$	$\|\|g(f(x^i)) - x^i\|\|^2$
1	0.5	0.75	[-0.625, -0.125]	0.406
2	2.3	2.15	[-0.925, -1.225]	2.356
3	3.1	3.05	[-1.475, -1.575]	4.656
4	3.9	3.95	[-2.025, -1.925]	7.8

12. (1 point)

Answer: A
For (a,b)=(1,1)
Loss=$\frac{0.0010+0.0171+0.0018+0.0017}{4} = 0.0054$

x	y	x+1	x+1-y	$(x+y-1)^2$
-1	0.0319	0	-0.0319	0.0010
0	0.8692	1	0.1308	0.0171
1	1.9566	2	0.0434	0.0018
2	3.0343	3	-0.0343	0.0017

Similarly,
for (a,b)=(1,2)
Loss= 1.0592
for (a,b)=(2,1)
Loss=1.5086
for (a,b)=(2,2)
Loss=3.562

13. (1 point)

Answer: g: 2.964 [Range could be 2 to 4]
h: 11.924 [Range could be 11 to 13]

$Loss_g = \frac{1.44+2.89+0.49+1+9}{5} = 2.964$
$Loss_h = \frac{0.04+0.09+18.49+25+16}{5} = 11.924$

$\mathbf{X}$	y	$g(\mathbf{X})$	$h(\mathbf{X})$	$y - g(\mathbf{X})$	$y - h(\mathbf{X})$	$(y - g(\mathbf{X}))^2$	$(y - h(\mathbf{X}))^2$
[2]	5.8	7	6	-1.2	-0.2	1.44	0.04
[3]	8.3	10	8	-1.7	0.3	2.89	0.09
[6]	18.3	19	14	-0.7	4.3	0.49	18.49
[7]	21	22	16	-1	5	1	25
[8]	22	25	18	-3	4	9	16

14. (2 points)

Answer: g: 1/6
h: 1/2

$\mathbf{X}$	y	g(X)	h(X)
[4, 2]	+1	1	-1
[8, 4]	+1	1	1
[2, 6]	-1	-1	-1
[4, 10]	-1	-1	1
[10, 2]	+1	1	1
[12, 8]	-1	1	1

15. (1 point)

Answer: Loss=34.5

$\mathbf{X}$	f(X)	g(f(X))	g(f(X))-X	$(\parallel g(f(\mathbf{X})) - \mathbf{X} \parallel)^2$
[1,2,3]	2.5	[2.5,5,7.5]	[1.5,3,4.5]	31.5
[2,3,4]	4	[4,8,12]	[2,5,8]	93
[-1,0,1]	-0.5	[-0.5,-1,-1.5]	[0.5, -1 , -2.5]	7.5
[0,1,1]	1	[1,2,3]	[1,1,2]	6

$$Loss = \frac{31.5+93+7.5+6}{4} = 34.5$$

Course: Machine Learning - Foundations
Week 2 - Test Questions

1. (2 points)

   Answer: D
   Option A and B are discontinuous at x=1.Option C is discontinuous at x=2. So correct option is D

2. (1 point)

   Answer: D
   A, B, C are equivalent statements.

3. (2 points)

   Answer: B, D
   Right hand limit and left hand limit at x=3 are different so function is not continuous at x=3

4. (1 point)

   Answer: 1.011
   Linear approximation of e^x is 1+x. Here x=0.011

5. (1 point)

   Answer: 1.975
   Linear approximation of $\sqrt{x}$ is,
   $\sqrt{x} - (x - x^*)\frac{1}{2\sqrt{x}}$.
   Here x= 4 and $x^* = 3.9$ $\sqrt{3}.9 = \sqrt{4} - (4 - 3.9)\frac{1}{2\sqrt{4}}$=1.975

6. (1 point)

   Answer: Two vectors are perpendicular if their inner product is zero

7. (2 points)

   Answer: B
   $f(x) \approx f(v) + \nabla f(v)^T(x - v)$
   $= 16 + [12 \quad 12]\begin{bmatrix} x-2 \\ y-2 \end{bmatrix}$
   $= 12x + 12y - 32$

8. (1 point)

   Answer: A
   $[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}]$
   $= [3x^2y^2 \quad 2x^3y]$
   $= [12 \quad 4]$

9. (1 point)

   Answer: C
   $[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}]$
   $= [3x^2 \quad 2y \quad 3z^2]$
   $= [0 \quad 2 \quad 3]$

10. (1 point)

    Answer: As per Cauchy Shwarz inequality,
    $- \| a \| \| b \| \leq a.b \leq \| a \| \| b$

11. (2 points)

    Answer: 0.816
    $\nabla f = [3x^2 \quad 2y \quad 3z^2]^T$
    at (1,1,1)
    $\nabla f = [3 \quad 2 \quad 3]^T$
    $\| i - 2j + k \| = \sqrt{6}$
    Directional derivative=$\frac{3\times 1 - 2\times 2 + 3\times 1}{\sqrt{6}}$

12. (2 points)

    Answer: A
    $\nabla f = [2 \quad 3y^2 \quad 4z]^T$
    at (1,0,1)
    $\nabla f = [2 \quad 0 \quad 4]^T$
    $\| [2 \quad 0 \quad 4]^T \| = \sqrt{20}$
    direction of steepest ascent=$\nabla f / \| f \|$

13. (2 points)

    Answer: 0.577
    $\nabla f = [1 \quad 1 \quad 1]^T$
    at (-1,1,0)
    $\nabla f = [1 \quad 1 \quad 1]^T$
    $\| i - j + k \| = \sqrt{3}$
    Directional derivative=$\frac{1\times 1 - 1\times 1 + 1\times 1}{\sqrt{3}}$

14. (1 point)

    Answer: Line through $u \in R^d$ along $v \in R^d$ is given by
    $x = u + \alpha v$
    where, $\alpha \in R$ and $x \in R^d$

Course: Machine Learning - Foundations
Week 3: Test questions

1. (1 point)

   Answer: $Length = \sqrt{1^2 + 2^2 + (-1)^2}$

2. (1 point)

   Answer: $1 \times -1 + 2 \times 1 + 3 \times 5 = 16$

3. (1 point)

   Answer: $rank + nullity = n$
   $or, 1 + nullity = 3$
   $or, nullity = 2$

4. (1 point)

   Answer: for option D inner product is zero

5. (1 point)

   Answer: Applying $R_2 - 2R_1$ and $R3 - 3R_1$ we get rank 1

6. (1 point)

   Answer: D

7. (1 point)

   Answer: $5Peaches + 6oranges = 150$
   $10Peaches + 12oranges = 300$
   in matrix form $\begin{bmatrix} 5 & 6 & 150 \\ 10 & 12 & 300 \end{bmatrix}$

8. (1 point)

   Answer: Solving $A^TA\hat{\theta} = A^Tb$ with $A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \\ 3 & 1 \end{bmatrix}$, $b = \begin{bmatrix} 6 \\ 3 \\ 15 \end{bmatrix}$ we get $\hat{\theta} = (3, 5)$

9. (1 point)

   Answer: Reducing the given matrix to echelon form we get $\begin{bmatrix} 1 & 0 & 9 & 2 \\ 0 & 1 & -3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$
   writing in equation form,
   $x_1 + 9x_3 + 2x_4 = 0$
   $x_2 - 3x_3 + x_4 = 0$
   Let $x_3 = s$ and $x_4 = t$ Therefore, $x_1 = -9s - 2t$

$x_2 = 3s - t$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -9s - 2t \\ 3s - t \\ s \\ t \end{bmatrix}$$

10. (1 point)

    Answer: Check dot product

11. (1 point)

    Answer: Use projection of vector $\vec{u}$ on $\vec{v}$ as $p = \dfrac{v^T u}{v^T v} v$

12. (1 point)

    Answer: A

13. (1 point)

    Answer: Use projection of vector $\vec{u}$ on $\vec{v}$ as $p = \dfrac{v^T u}{v^T v} v$

Course: Machine Learning - Foundations
Week 4: Test questions

1. (1 point)

   Answer: A

2. (1 point)

   Answer: The determinant, column space and rank of a matrix are not affected by row operations on the matrix

3. (1 point)

   Answer: D

4. (1 point)

   Answer: A,D

5. (1 point)

   Answer: Determinant= product of eigen values

6. (1 point)

   Answer: Trace=sum of eigen values

7. (1 point) If the eigenvalues of a matrix are -1, 0 and 4, then its trace and determinant are
   Trace:_____
   Determinant:_____

   Answer: Determinant= product of eigen values
   Trace=sum of eigen values

8. (1 point)

   Answer: To find characteristic polynomial obtain $|A - \lambda I|$

9. (1 point)

   Answer: Solve for λ, $|A - \lambda I| = 0$

10. (1 point)

    Answer: If λ is an eigen value of A then λ^n is eigen value of A^n and vice versa

11. (1 point)

    Answer: A

12. (1 point)

    Answer: B

13. (2 points)

    Answer: 0, 5

14. (2 points)

    Answer: solve $(A - \lambda I)x = 0$, where λ is the eigen value and x is corresponding eigen vector

15. (2 points)

    Answer: Let us consider $P^{-1}AP = B$, Here B is an upper triangular matrix, So the eigenvalues are same as principal diagonal elements. Now, the eigenvalues of B are the eigenvalues of A. So eigenvalues of A^2 are eigenvalues of B squared.

16. (2 points)

    Answer: $\begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1.3 & 1.69 \\ 1 & 4 & 16 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1.5 \\ 2 \end{bmatrix}$

Course: Machine Learning - Foundations
Week 5: Test questions

1. (1 point) Consider two non-zero vectors $x \in \mathbb{C}^n$ and $y \in \mathbb{C}^n$. Suppose the inner product between x and y obeys commutative property (i.e., $x \cdot y = y \cdot x$), it implies that

   A. y must be a conjugate transpose of x

   B. y is equal to x

   C. y must be orthogonal to x

   D. y must be a scalar (possibly complex) multiple of x

   Answer: C
   For orthogonal vectors dot product is zero.

2. (1 point) The inner product of two distinct vectors x and y that are drawn randomly from $\mathbb{C}^{100}$ is $0.8 - 0.37i$.The vector x is scaled by a scalar $1 - 2i$ to obtain a new vector z, then the inner product between z and y is

   A. $0.06 - 1.97i$

   B. $1.54 - 1.23i$

   C. $1.54 + 1.23i$

   D. $0.8 - 0.37i$

   E. Not possible to calculate

   Answer: C
   $x.y = \bar{x}^T y$
   $z = cx$
   $z.y = \bar{c}\bar{x}^T y$
   $z.y = \bar{c}x.y = 1.54 + 1.23i$

3. (1 point) Select the correct statement(s).The Eigen-value decomposition for the matrix $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

   A. doesn't exist over $\mathbb{R}$ but exists over $\mathbb{C}$

   B. doesn't exist over $\mathbb{C}$ but exists over $\mathbb{R}$

   C. neither exists over $\mathbb{R}$ nor exists over $\mathbb{C}$

   D. exists over both $\mathbb{C}$ and $\mathbb{R}$

   Answer: A
   Eigenvalues of A are complex.

4. (1 point) Consider the complex matrix $S = \begin{bmatrix} 1 & 1+i & -2-2i \\ 1-i & 1 & -i \\ -2+2i & i & 1 \end{bmatrix}$. The matrix is

   A. Hermitian and Symmetric

   B. Symmetric but not Hermitian

   C. Neithet Hermitian nor Symmetric

   D. Hermitian but not Symmetric

Answer: D
$S^T \neq S$
$S^* = S$

5. (1 point) Suppose that an unitary matrix U is multiplied by a diagonal matrix D with $d_{ii} \in \mathbb{R}$, then the resultant matrix will always be unitary. The statement is

   A. True

   B. False

Answer: A
Let $A = UD$
Where, U=unitary matrix
D=Diagonal matrix
$A^* = (UD)^*$
$AA^* = (UD)(UD)^*$
$AA^* = (UDD^*U^*)$

6. (3 points) The eigenvectors of matrix $A = \begin{bmatrix} 3 & 2-i & -3i \\ 2+i & 0 & 1-i \\ 3i & 1+i & 0 \end{bmatrix}$ are

   A. $\begin{bmatrix} -1 \\ 1+2i \\ 1 \end{bmatrix}, \begin{bmatrix} 1-21i \\ 6-9i \\ 13 \end{bmatrix}, \begin{bmatrix} 1+3i \\ -2-i \\ 5 \end{bmatrix}$

   B. $\begin{bmatrix} 1 \\ 1-2i \\ 1 \end{bmatrix}, \begin{bmatrix} 1-21i \\ 6-9i \\ 13 \end{bmatrix}, \begin{bmatrix} 1+3i \\ -2-i \\ 5 \end{bmatrix}$

   C. $\begin{bmatrix} -1 \\ 1-2i \\ -1 \end{bmatrix}, \begin{bmatrix} 1-21i \\ 6-9i \\ 13 \end{bmatrix}, \begin{bmatrix} 1+3i \\ -2-i \\ 5 \end{bmatrix}$

   D. $\begin{bmatrix} -1 \\ 1+2i \\ 1 \end{bmatrix}, \begin{bmatrix} 1-21i \\ 6-9i \\ 13 \end{bmatrix}, \begin{bmatrix} 1-3i \\ 2-i \\ -5 \end{bmatrix}$

Answer: A
$|A - \lambda I| = 0$
$\lambda^3 - 3\lambda^2 - 16\lambda - 12 = 0$
$\lambda = -1, -2, 6$

For $\lambda = -1$
Eigenvector, $v_1 = \begin{bmatrix} -1 \\ 1+2i \\ 1 \end{bmatrix}$
For $\lambda = -2$
Eigenvector, $v_2 = \begin{bmatrix} 1-21i \\ 6-9i \\ 13 \end{bmatrix}$
For $\lambda = 6$
Eigenvector, $v_3 = \begin{bmatrix} 1+3i \\ -2-i \\ 5 \end{bmatrix}$

7. (1 point) A matrix $A = \frac{1}{2}\begin{bmatrix} k+i & \sqrt{2} \\ k-i & \sqrt{2}i \end{bmatrix}$ is unitary if k is
   A. $\frac{1}{2}$
   B. 1
   C. -$\frac{1}{2}$
   D. -1
   E. ± 1
   F. $\pm\frac{1}{2}$

Answer: B
$\dfrac{1}{2}\begin{bmatrix} k+i & \sqrt{2} \\ k-i & \sqrt{2}i \end{bmatrix} \dfrac{1}{2}\begin{bmatrix} k-i & k+i \\ \sqrt{2} & -\sqrt{2}i \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
$\dfrac{1}{4}\begin{bmatrix} k^2+3 & k^2-1+2i(k-1) \\ k^2-1-2i(k-1) & k^2+3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
$k = 1$

8. (3 points) A matrix $A = \begin{bmatrix} 1 & 1+i \\ 1-i & 1 \end{bmatrix}$ can be written as $A = UDU^*$, where U is a unitary matrix and D is a diagonal matrix. Then, U and D, respectively, are
   A. $U = \begin{bmatrix} 1+i & \sqrt{2} \\ \sqrt{2} & 1-i \end{bmatrix}$, $D = \begin{bmatrix} 1+\sqrt{2} & 0 \\ 0 & 1-\sqrt{2} \end{bmatrix}$
   B. $U = \begin{bmatrix} -1+i & \sqrt{2} \\ \sqrt{2} & -1-i \end{bmatrix}$, $D = \begin{bmatrix} 1+\sqrt{2} & 0 \\ 0 & 1-\sqrt{2} \end{bmatrix}$

C. $U = \begin{bmatrix} 1+i & \sqrt{2} \\ \sqrt{2} & 1-i \end{bmatrix}$, $D = \begin{bmatrix} -1+\sqrt{2} & 0 \\ 0 & 1-\sqrt{2} \end{bmatrix}$

D. $U = \begin{bmatrix} 1-i & \sqrt{2} \\ \sqrt{-2} & 1-i \end{bmatrix}$, $D = \begin{bmatrix} 1+\sqrt{2} & 0 \\ 0 & 1-\sqrt{2} \end{bmatrix}$

Answer: A

To find eigenvalues, $|A - \lambda I| = 0$

$\lambda^2 - 5\lambda + 4 = 0$

$\lambda = 1+\sqrt{2}, 1-\sqrt{2}$

Find eigenvectors

For $\lambda = 1+\sqrt{2}$,

$v_1 = \begin{bmatrix} 1+i \\ \sqrt{2} \end{bmatrix}$

For $\lambda = 1-\sqrt{2}$,

$v_2 = \begin{bmatrix} \sqrt{2} \\ 1-i \end{bmatrix}$

$U = [v_1 \quad v_2]$

$D = \begin{bmatrix} 1+\sqrt{2} & 0 \\ 0 & 1-\sqrt{2} \end{bmatrix}$

9. (2 points) The matrix $Z = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$ has

A. only real eigenvalues.

B. two real and two complex eigenvalue.

C. three real and one complex eigenvalues.

D. all complex eigenvalues

Answer: B

$\lambda = 1, -1, i, -i$

10. (1 point) (Multiple select) Which of the following matrices is/are unitary?

A. $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$

B. $\begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

C. $\begin{bmatrix} -\cos\theta & \sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

D. $\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$

E. $\begin{bmatrix} -\cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$

Answer: C, D
Check $UU^* = I$

11. (2 points) Let U and V be two unitary matrices. Then

 1. UV is unitary.
 2. $U + V$ is unitary.

 A. Both statements are true.
 B. Both statements are false.
 C. 1. is false.
 D. 2. is false.

Answer: D
Addition of two unitary matrices may not be unitary.

12. (2 points) (Multiple select) Which of the following is/are eigenvectors of the Hermitian matrix $A = \begin{bmatrix} 1 & 1+i \\ 1-i & 2 \end{bmatrix}$

 A. $\begin{bmatrix} -1-i \\ 1 \end{bmatrix}$
 B. $\begin{bmatrix} -2-2i \\ 2 \end{bmatrix}$
 C. $\begin{bmatrix} \frac{1+i}{2} \\ 1 \end{bmatrix}$
 D. $\begin{bmatrix} 1+i \\ 2 \end{bmatrix}$
 E. All of these.

Answer: E
To find eigenvalues, $|A - \lambda I| = 0$
$\lambda^2 - 3\lambda = 0$
$\lambda = 0, 3$

For $\lambda = 0$
Eigenvector, $v_1 = \begin{bmatrix} 1+i \\ -1 \end{bmatrix}$
For $\lambda = 3$
Eigenvector, $v_2 = \begin{bmatrix} 1+i \\ 2 \end{bmatrix}$

Week 6 Graded Assignment Solution

1.

$$f(x) = x^2 + y^2$$
$$f_x = \frac{\partial(x,y)}{\partial x} = 2x$$
$$f_y = \frac{\partial(x,y)}{\partial y} = 2x$$

Put $f_x = 0$ and $f_y = 0$,

$$f_x = 0 \Rightarrow 2x = 0 \Rightarrow x = 0$$

$$f_y = 0 \Rightarrow 2y = 0 \Rightarrow y = 0$$

Thus, stationary points $= (0, 0)$

2.

$$A = \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix} \Leftrightarrow \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

Here, $a = 4 > 0$ and $ac - b^2 = 4(2) - 2^2 = 4 > 0$
So, A is positive definite.

Now,

$$B = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \Leftrightarrow \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

Here, $a = 1 > 0$ and $ac - b^2 = 1(2) - 1^1 = 1 > 0$
So, B is positive definite.

$A + B = \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 3 \\ 3 & 4 \end{bmatrix} \Leftrightarrow \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ $a = 5 > 0$ and $ac - b^2 = 5(4) - 3^2 = 11 > 0$
So, $A + B$ is also positive definite.

3.

$$A = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}$$

The characteristic polynomial is

$$\lambda^3 - 6\lambda^2 + \{3 + 3 + 3\}\lambda - 4 = 0$$

FORMULA:
$x^3 - [trace(A)]x^2 + \Sigma[\text{Minors of diagonal elements(A)}]x - det(A) = 0$

$$\lambda^3 - 6\lambda^2 + 9\lambda - 4 = 0$$

Solving we get $\lambda = 4, 1, 1$ Since all eigenvalues are greater than zero, A is positive definite.

4.
$$f(x,y) = 2x^2 + 2xy + 2y^2 - 6x$$
$$f_x = 4x + 2y - 6$$
$$f_y = 2x + 4y$$

For stationary point, $f_x = 0$ and $f_y = 0$

$$4x + 2y - 6 = 0 \Rightarrow 2x + y = 3$$
$$2x + 4y = 0 \Rightarrow x = -2y$$

Solving we get $x = 2$ and $y = -1$

Hence, the stationary point $= (2, -1)$

5.
$$\begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$
$$\begin{bmatrix} ax + dy + gz & bx + ey + hz & cx + fy + iz \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$
$$ax^2 + ey^2 + iz^2 + (b + d)xy + (g + c)xz + (f + h)yz$$

Compare this with $x^2 + y^2 - z^2 - xy + yz + xz$ we get
$a = 1, e = 1, i = -1, b + d = -1, c + g = 1, f + h = 1$

Only in options A and C, the diagonal elements are $1, 1, -1$

- If we check in option C, then $f + h = -1$ which is not satisfied. So, this is incorrect.
- If we check in option A, then all $b + d = -1, c + g = 1, f + h = 1$ are satisfied.

So, option A is correct.

6.

$$f(x,y) = 3x^2 + 4xy + 2y^2$$
$$f_x = 6x + 4y \Rightarrow f_{xx} = 6$$
$$f_y = 4x + 4y \Rightarrow f_{yy} = 4$$

Since $f_{xx} > 0$ and $f_{yy} > 0$, the point $(0,0)$ is a minima.

7.

$$A = \begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix} \Leftrightarrow \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

Here, $a = 4 > 0$ and $ac - b^2 = 4(3) - 2^2 = 8 > 0$
So, A is positive definite.

8.

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \Leftrightarrow \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

Here, $a = 1 > 0$ and $ac - b^2 = 1(1) - 2 * 2 = -1 < 0$
Since $a > 0$ but $ac - b^2 < 0$, A is NOT positive definite.

9.

$$A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 7 \end{bmatrix}$$

Since A is a diagonal matrix, the eigenvalues are $3, 5$ and 7.

Now since all eigenvalues are greater than 0, A is positive definite.

10.

$$A = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}$$

$$AA^T = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 1 & 2 \\ 1 & 1 & 0 \\ 2 & 0 & 2 \end{bmatrix}$$

The characteristics polynomial of AA^T:

$$\lambda^3 - 6\lambda^2 + 6\lambda = 0$$

Solving we get

$$\lambda = 0, 3 \pm \sqrt{3}$$

Now, $\sigma = \sqrt{\lambda}$
So,

$$\sigma_1 = \sqrt{3+\sqrt{3}}\, and\, \sigma_2 = \sqrt{3-\sqrt{3}}$$

are the non-zero singular values.

11.

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

$$AA^T = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

The characteristics polynomial of AA^T:

$$\lambda^2 - 4\lambda + 4 = 0$$

Solving we get

$$\lambda = 2, 2$$

$$\sigma = \sqrt{\lambda} = \sqrt{2}$$

$$\Sigma = \begin{bmatrix} \sqrt{2} & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 \end{bmatrix}$$

Now for $\lambda = 2$,

$$(AA^T - 2I)X = 0$$

$$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} - 2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0$$

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0$$

$$v = k \times \begin{bmatrix} x \\ y \end{bmatrix}$$

for $x = 1$, $y = 0$; $v_1 = k \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

for $x = 0$, $y = 1$; $v_2 = k \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

$$Q_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$y_1 = \frac{A^T x_1}{\sigma_1}$$

$$y_1 = \frac{1}{\sqrt{2}} \times \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$$

Similarly,

$$y_2 = \frac{1}{\sqrt{2}} \times \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

For the other two

$$\begin{bmatrix} a & b & c & d \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} = 0$$

and

$$\begin{bmatrix} a & b & c & d \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = 0$$

From above we have

$$a + c = 0 \, and \, b + d = 0$$

Let $c = k_1$ and $d = k_2$

$$k_1 \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + k_2 \begin{bmatrix} - \\ -1 \\ 0 \\ 1 \end{bmatrix} Normalizing \rightarrow \frac{1}{\sqrt{2}} k_1 \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} ; \frac{1}{\sqrt{2}} k_2 \begin{bmatrix} - \\ -1 \\ 0 \\ 1 \end{bmatrix} Now,$$

$$y_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} for \, k_1 = -1 \, and \, y_4 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix} for \, k_2 = -1$$

So,

$$Q_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}$$

$$Q_2^T = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}$$

1. Consider the following matrix:

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 1 \end{bmatrix}$$

Which of the following vectors is not an eigenvector of this matrix ?

A. $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

B. $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$

C. $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$

D. $\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$

Solution: Option D is the correct answer. For each of the vectors given in the options you can compute the product Ax. For example, consider $x = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

$$Ax = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 4 \end{bmatrix} = 4 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 4x$$

Hence, $x = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ is an eigenvector of A. Similarly, you can show that $x = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ and $x = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ are also eigenvectors of this matrix with the corresponding eigen values being 1 and -1 respectively. You can then also check that $x = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ is not an eigenvector of this matrix.

2. Consider a square matrix $A \in \mathbb{R}^{3\times 3}$ such that $A^T = A$. My friend told me that the following three vectors are the eigenvectors of this matrix A:

$$x = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, y = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, z = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

Is my friend telling the truth ?

A. Yes

B. No

C. Can't say without knowing all the elements of A

D. Yes, only if all the diagonal elements of A are 1

Solution: Note that A is a square symmetric matrix ($\because A \in \mathbb{R}^{3\times 3}$ and $A^T = A$). We know that the eigenvectors of a square symmetric matrix are orthogonal. In other words, if x, y, z are the eigenvectors of A then $x^T y = x^T z = y^T z = 0$. You can easily verify that this is not the case. Hence, my friend is not telling the truth. **Option B** is the correct answer.

A. A is positive definite.

B. A is positive semi-definite.

C. A is neither positive definite nor positive semi-definite.

D. Can not be determined

Answer: A

This is a diagonal matrix. For a diagonal matrix, the eigenvalues are elements on its principal diagonal. Since all the eigenvalues (diagonal elements) are positive, the matrix is a positive definite matrix.

Questions 10-15 are based on common data

Consider these data points to answer the following questions:

$$x_1 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, x_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, x_3 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$$

10. (1 point) The mean vector of the data points x_1, x_2, x_3 is

A. $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

B. $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

C. $\begin{bmatrix} 0.9 \\ 0.6 \\ 0.3 \end{bmatrix}$

D. $\begin{bmatrix} 0.5 \\ 0.5 \\ 0.5 \end{bmatrix}$

Answer: B

$$\text{Mean vector} = \bar{X} = \Sigma_{i=1}^{n} \frac{1}{n} x_i = \frac{1}{3} \begin{bmatrix} (0+1+2) \\ (1+1+1) \\ (2+1+0) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

11. (2 points) The covariance matrix $C = \frac{1}{n} \Sigma_{i=1}^{n} (x_i - \bar{x})(x_i - \bar{x})^T$ of the data points x_1, x_2, x_3 is

A. $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

B. $\begin{bmatrix} 0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 \end{bmatrix}$

C. $\begin{bmatrix} 0.7 & 0 & -0.7 \\ 0 & 0 & 0 \\ -0.7 & 0 & 0.7 \end{bmatrix}$

D. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Answer: C

$$C = \frac{1}{3}(\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}) = \frac{1}{3} \begin{bmatrix} 2 & 0 & -2 \\ 0 & 0 & 0 \\ -2 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 0.7 & 0 & -0.7 \\ 0 & 0 & 0 \\ -0.7 & 0 & 0.7 \end{bmatrix}$$

12. (2 points) The eigenvalues of the covariance matrix $C = \frac{1}{n}\Sigma_{i=1}^{n}(x_i - \bar{x})(x_i - \bar{x})^T$ are

A. 2, 0, 0

B. 1, 1, 1

C. 1.4, 0, 0

D. 0.5, 0, 0.5

Answer: C

Characteristics equation:

$$\begin{bmatrix} \frac{7}{10} - \lambda & 0 & -\frac{7}{10} \\ 0 & -\lambda & 0 \\ -\frac{7}{10} & 0 & \frac{7}{10} - \lambda \end{bmatrix} I = 0$$

The determinant of the obtained matrix is $\lambda^2(\frac{7}{5} - \lambda) = 0$

Eigenvalues:

The roots are $\lambda_1 = \frac{7}{5} = 1.4, \lambda_2 = 0, \lambda_3 = 0$

Eigenvectors:

$$\lambda_1 = 1.4, \begin{bmatrix} \frac{7}{10} - \lambda & 0 & -\frac{7}{10} \\ 0 & -\lambda & 0 \\ -\frac{7}{10} & 0 & \frac{7}{10} - \lambda \end{bmatrix} = \begin{bmatrix} -\frac{7}{10} & 0 & -\frac{7}{10} \\ 0 & -\frac{7}{5} & 0 \\ -\frac{7}{10} & 0 & -\frac{7}{10} \end{bmatrix}$$

The null space of this matrix is $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, Corresponding eigenvector is, $u_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -0.7 \\ 0 \\ 0.7 \end{bmatrix}$

$$\lambda_2 = \lambda_3 = 0, \begin{bmatrix} \frac{7}{10} - \lambda & 0 & -\frac{7}{10} \\ 0 & -\lambda & 0 \\ -\frac{7}{10} & 0 & \frac{7}{10} - \lambda \end{bmatrix} = \begin{bmatrix} \frac{7}{10} & 0 & -\frac{7}{10} \\ 0 & 0 & 0 \\ -\frac{7}{10} & 0 & \frac{7}{10} \end{bmatrix}$$

The null space of this matrix are $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, and $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ Corresponding eigenvector are, $u_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $u_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.7 \\ 0 \\ 0.7 \end{bmatrix}$

13. (2 points) The eigenvectors of the covariance matrix $C = \frac{1}{n}\Sigma_{i=1}^{n}(x_i - \bar{x})(x_i - \bar{x})^T$ are (Note: The eigenvectors should be arranged in the descending order of eigenvalues from left to right in the matrix.)

    A. $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$

    B. $\begin{bmatrix} 0.5 & 0 & 0.5 \\ 0 & 0 & 0 \\ 0.5 & 0 & 0.5 \end{bmatrix}$

    C. $\begin{bmatrix} -0.7 & 0 & 0.7 \\ 0 & 1 & 0 \\ 0.7 & 0 & 0.7 \end{bmatrix}$

    D. $\begin{bmatrix} -1 & 0 & 0 \\ -1 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}$

    Answer: C

    Refer to the solution of the previous question.

14. (2 points) The data points x_1, x_2, x_3 are projected onto the one dimensional space using PCA as points z_1, z_2, z_3 respectively. (Use eigenvector with the maximum eigenvalue for this projection.)

    A. $z_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $z_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $z_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

B. $z_1 = \begin{bmatrix} 0.5 \\ 0.5 \\ 0.5 \end{bmatrix}$, $z_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, $z_3 = \begin{bmatrix} -0.5 \\ -0.5 \\ -0.5 \end{bmatrix}$

C. $z_1 = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}$, $z_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $z_3 = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}$

D. $z_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, $z_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, $z_3 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$

Answer: D

$\lambda_1 = 1.4$, $u_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$

$$z_1 = \frac{1}{\sqrt{2}}\left(\begin{bmatrix} 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\right)\frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

$$z_2 = \frac{1}{\sqrt{2}}\left(\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\right)\frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$z_3 = \frac{1}{\sqrt{2}}\left(\begin{bmatrix} 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\right)\frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

15. (1 point) The approximation error J on the given data set is given by $\Sigma_{i=1}^{n} ||x_i - z_i||^2$. What is the reconstruction error?

A. 3

B. 5

C. 10

D. 20

Answer: A

Approximation Error, $J = \frac{1}{n}\Sigma_{i=1}^{n} ||x_i - z_i||^2 = \frac{1}{3}[(1^2 + 1^2 + 1^2) + (1^2 + 1^2 + 1^2) + (1^2 + 1^2 + 1^2)] = 3$

Course: Machine Learning - Foundations
Week [redacted]: Test questions

1. (2 points) Two positive numbers have a sum of 60. What is the minimum product of one number times the square of other number?

   A. 0

   B. 900

   C. 60

   D. 240

   Answer: A
   Let the two numbers be x and y
   x+y=60
   objective function from the question will be,
   $f(x) = x^2(60 - x)$
   For optima $f'(x) = 0$, $120x - 3x^2 = 0$
   $x = 0, 40$
   Product is minimum when x=0.

2. (2 points) (Multiple select) The point on $y = x^2 + 1$ closest to (0,2) is

   A. (0.707, 1.5)

   B. (0.707, -1.5)

   C. (-0.707,1.5)

   D. (-0.707, -1.5)

   Answer: A,C
   Objective function $f(x) = (x - 0)^2 + (x^2 + 1 - 2)^2$
   $f(x) = x^4 - x^2 + 1$
   For minima $f'(x) = 0$
   $4x^3 - 2x = 0$
   $x = 0, 0.707, -0.707$
   Corresponding $y = 1, 1.5, 1.5$

3. (2 points) The volume of the largest cone that can be inscribed in a circle of radius 6 m is (correct up to two decimal places)

   _____

   Answer: 268.19 m^3
   $V = \frac{1}{3}\pi r^2 h$
   $r = \sqrt{36 - x^2}$
   $h = 6 + x$

For maxima, $V'(x) = 0$
$-3x^2 - 12x + 36 = 0$
$x = 2, -6$
x can not be nagative.
So $r = 5.65$
$h = 8$
$V = 268.19$

(Questions 5-8 have common data) A firm produces two products A and B. Maximum production capacity is 500 for total production. At least 200 units must be produced every day. Machine hours consumption per unit is 5 hours for A and 3 hours for B. At least 1000 machine hours must be used daily. Manufacturing cost is Rs 30 for A and Rs 20 for B.
Let x_1 = No of units of A produced per day
and x_2 = No of units of B produced per day

4. (1 point) The objective function for above problem is

   A. min $f(x) = 30x_1 + 20x_2$
   B. min $f(x) = 15x_1 + 55x_2$
   C. min $f(x) = 5x_1 + 155x_2$
   D. min $f(x) = 30x_1 - 20x_2$

   Answer: A
   We should minimise cost function.
   Objective function is
   min $f(x) = 30x_1 + 20x_2$

5. (2 points) The constraint due to maximum production capacity is

   A. $x_1 + x_2 \geq 500$
   B. $x_1 + x_2 \leq 500$
   C. $x_1 + x_2 \neq 500$
   D. $x_1 + x_2 = 500$

   Answer: B
   Maximum production capacity is 500.

6. (2 points) The constraint due to minimum production capacity is

   A. $x_1 + x_2 = 200$
   B. $x_1 + x_2 \leq 200$
   C. $x_1 + x_2 \geq 200$
   D. $x_1 + x_2 \neq 200$

   Answer: C
   Minimum production capacity is 200.

7. (2 points) The constraint due to machine hour consumption is

   A. $5x_1 + 3x_2 \leq 1000$

   B. $5x_1 + 3x_2 \neq 1000$

   C. $5x_1 + 3x_2 = 1000$

   D. $5x_1 + 3x_2 \geq 1000$

   Answer: D

   1000 machine hours must be used daily.

   (Questions 9-11 have common data)
   A factory manufactures two products A and B. To manufacture one unit of A, 1 machine hours and 2 labour hours are required. To manufacture product B, 2 machine hours and 1 labour hours are required. In a month, 200 machine hours and 140 labour hours are available. Profit per unit for A is Rs. 45 and for B is Rs. 35.
   Let x_1=Number of units of A produced per month
   and x_2=Number of units of B produced per month

8. (1 point) The objective function for above problem is

   A. $\max f(x) = 45x_1 + 35x_2$

   B. $\min f(x) = 45x_1 + 35x_2$

   C. $\max f(x) = 35x_1 + 45x_2$

   D. $\min f(x) = 35x_1 + 45x_2$

   Answer: A

   We need to maximize profit.

9. (2 points) The constraint for machine hours is

   A. $x_1 + 2x_2 \geq 200$

   B. $x_1 + 2x_2 \leq 200$

   C. $x_1 + 2x_2 \neq 200$

   D. $x_1 + 2x_2 = 200$

   Answer: B

   Total machine hours available=200.

10. (2 points) The constraint for labour hours is

    A. $2x_1 + x_2 = 140$

    B. $2x_1 + x_2 \leq 140$

    C. $2x_1 + x_2 \geq 140$

    D. $2x_1 + x_2 \neq 140$

    Answer: B

    Total labour hour available is 140.

11. (2 points) (Multiple select) Gradient of a continuous and differentiable function
   A. is zero at a minimum
   B. is non zero at a maximum
   C. is zero at a saddle point
   D. decreases as you get closer to minimum

   Answer: A,C,D
   For critical points gradient of a function is 0.
   As we move towards minima gradient decreases.

12. The value of a function at point 10 is 100. The values of the function's first and second order derivatives at this point are 20 and 2 respectively. What will be the function's approximate value correct up to two decimal places at the point 10.5 (Use second order approximation)?

   Answer: 110.25
   According to Taylor's series,
   $f(x+h) = f(x) + hf'(x) + \frac{h^2 f'(x)}{2} + ...$
   Here $x = 10, h = 0.5$
   $\therefore f(x+h) = 110.25$

13. (2 points) For the function $f(x) = x\sin(x) - 1$, with an initial guess of $x_0 = 2.5$, and step size of 0.1, as per gradient descent algorithm, what will be the value of the function after 4 iterations? (Correct up to 3 decimal places)

   Answer: -1.710 (-1.624 to-1.795)
   $x_{n+1} = x_n - \eta f'(x)$
   After first iteration $x_1 = 2.64$
   After second iteration $x_2 = 2.823$
   After third iteration $x_3 = 3.059$
   After fourth iteration $x_4 = 3.355$
   $f(3.355) = -1.710$

14. (2 points) The value of $f(x_1, x_2) = 4x_1^2 - 4x_1x_2 + 2x_2^2$ with an initial guess of (2, 3) after two iterations of gradient descent algorithm will be .... Take the step size $\eta = \frac{1}{t+1}$, where t= 0,1,2....

   Answer: 130
   $x_{n+1} = x_n - \eta\nabla f(x)$
   $\nabla f = \begin{bmatrix} 8x_1 - 4x_2 \\ -4x_1 + 4x_2 \end{bmatrix}$
   $x_1 = \begin{bmatrix} -2 \\ -1 \end{bmatrix}$

$x_2 = \begin{bmatrix} 4 \\ -3 \end{bmatrix}$
$f(4, -3) = 130$

15. (2 points) The point of minimum for the function $f(x_1, x_2) = x_1^2 - x_1x_2 + 2x_2^2$ with an initial guess of (3, 2) with step size=0.5 using gradient descent algorithm after second iteration will be .... (correct up to 3 decimal places)

Answer: 2.312 (2.196 to 2.428)
$x_{n+1} = x_n - \eta\nabla f(x)$
$\nabla f = \begin{bmatrix} 2x_1 - x_2 \\ -x_1 + 4x_2 \end{bmatrix}$
$x_1 = \begin{bmatrix} 1 \\ -0.5 \end{bmatrix}$
$x_2 = \begin{bmatrix} -0.25 \\ 1 \end{bmatrix}$
$f(-0.25, 1) = 2.312$

16. (2 points) Suppose we have n data points randomly distributed in space given by $D = \{x_1, x_2, ..., x_n\}$. A function $f(p)$ is defined to calculate the sum of distances of data points from a fixed point, say p. Let $f(p) = \sum_{i=1}^{n}(p - x_i)^2$. What is the value of p so that f(p) is minimum?

A. $x_1 + x_2 + ... + x_n$

B. $x_1 - x_2 + x_3 - x_4....$

C. $\dfrac{x_1 + x_2 + ... + x_n}{n}$

D. $\dfrac{x_1 - x_2 + x_3 - x_4....}{n}$

Answer: C
$f(p) = \sum_{i=1}^{n}(p - x_i)^2$
$f(p) = (p - x_1)^2 + ... + (p - x_n)^2$
$f'(p) = 2p(p - x_1) + ... + 2p(p - x_n)$
For minima $f'(p) = 0$
$(p - x_1) + (p - x_2) + ... + (p - x_n) = 0$
$np - (x_1 + x_2 + ... + x_n) = 0$
$p = \dfrac{x_1 + x_2 + ... + x_n}{n}$

Course: Machine Learning - Foundations
Graded Questions - Solution
Lecture Details: Week 8

1. (1 point) Points $(0,0),(5,0),(5,5),(5,0)$ forms a convex hull. Which of the following points are the part of this convex hull?

    A. (1,1)

    B. (1,-1)

    C. (-1,1)

    D. (-1,-1)

    Answer: A

    Let $(x_1,y_1)=(0,0),(x_2,y_2)=(5,0),(x_3,y_3)=(5,5),(x_4,y_4)=(5,0)$ Any point on the convex hull of these points will be the the set S such that

    $\{S=(x,y) \mid x=\lambda_1x_1+\lambda_2x_2+\lambda_3x_3+\lambda_4x_4, y=\lambda_1y_1+\lambda_2y_2+\lambda_3y_3+\lambda_4y_4, \lambda_1,\lambda_2,\lambda_3,\lambda_4 \in [0,1]\}$

    The convex hull of these 4 points forms a square and any point inside or on the square will be the part of the convex hull.

    The point (1,1) lies inside this square, while (1,-1), (-1,1) and (-1,=1) lies outside this square.

2. (1 point) Given S is a convex set and the points $x_1,x_2,x_3,x_4 \in S$. Which of the following points must be the part of convex set S:

    A. $0.1x_1+0.2x_2+0.3x_3+0.4x_4$

    B. $-0.1x_1+-0.2x_2+0.6x_3+0.7x_4$

    C. $0.1x_1+0.1^2x_2+0.1^3x_3+0.1^3x_4$

    D. $0.25x_1+0.25x_2+0.25x_3+0.25x_4$

    Answer: A, D

    The convex combination of the points will always be the part of the convex set and is known as convex hull.

    For the convex combination, the coefficients should be non-negative and should sum to 1. Therefore , options (A) and (D) are correct.

3. (1 point) Which of the following is a convex function in $\mathbb{R}^2$?

    A. $f(x)=x^2+y^2$

    B. $f(x)=-x^2-y^2$

    C. $f(x)=x^2-y^2$

D. None of these

Answer: A

$f(x) = x^2 + y^2 = v^T A v = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$

where $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, v = \begin{bmatrix} x \\ y \end{bmatrix}$

Here, $a > 0, ac - b^2 = 1.1 - 0^2 = 1 > 0$, This shows the matrix A is a positive definite matrix. Therefore, the function is a convex function.

4. (1 point) What is the boundary value of x so that the function $(x-3)^3 + (y+1)^2$ to remain convex?

   A. $x \geq 1$

   B. $x \geq 2$

   C. $x \geq 3$

   D. None of these

Answer: C

$f(x, y) = (x-3)^3 + (y+1)^2$

First order partial derivatives, $f_x = 3(x-3)^2, f_y = 2(y+1)$

Second order partial derivatives, $f_{xx} = 6(x-3), f_{xy} = 0, f_{yy} = 2$

The Second order partial derivatives with respect to x, $f_{xx} = 6(x-3)$ changes sign at the point where $x = 3$.

When $x \geq 3, f_{xx} \geq 0$ otherwise $f_{xx} < 0$

The hessian matrix will be,

$D = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix} = \begin{bmatrix} 6(x-3) & 0 \\ 0 & 2 \end{bmatrix} = 12(x-3)$

D will be non negative when $x \geq 3$ and the function remains convex.

5. (1 point) Select the most appropriate linear approximation of $f(x, y) = x^2 + y^2$ at a small step 0.02 in the direction (1,2) from the point (3,2)

   A. 11.02

   B. 10.28

   C. 9.98

   D. None of these

Answer: A

$\epsilon = 0.02, d = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \nabla f(x,y)_{(3,2)} = \begin{bmatrix} 2x \\ 2y \end{bmatrix}_{(3,2)} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$

The linear approximation of a function f at the point $(x + \epsilon d)$ is

$f(x,y) + \epsilon\ d^T \nabla f(x,y) = (3^2 + 1^2) + 0.02 * \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 6 \\ 4 \end{bmatrix} = 10 + 0.02 * 14 = 10.28$

6. (1 point) The minimum value of $f(x,y) = x^2 + 4y^2 - 2x + 8y$ subject to the constraint $x + 2y = 7$ is ______.

Answer: 27.00, Range 26.50 to 27.50

Given $f(x,y) = x^2 + 4y^2 - 2x + 8y,\ g(x,y) = x + 2y = 7$

$\nabla f(x,y) = \begin{bmatrix} 2x - 2 \\ 8y + 8 \end{bmatrix}, \nabla g(x,y) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

We find the values of x, y, λ that simultaneously satisfy the equations to get the extreme points $\nabla f(x,y) = \lambda \nabla g(x,y)$ and, $g(x,y) = x + 2y = 7$

Solving, $\nabla f(x,y) = \lambda \nabla g(x,y)$

$\implies \begin{bmatrix} 2x - 2 \\ 8y + 8 \end{bmatrix} = \lambda \begin{bmatrix} 1 \\ 2 \end{bmatrix} \implies x = (\lambda + 2)/2, y = (\lambda - 4)/4$

Solving, $g(x,y) = x + 2y = 7 \implies (\lambda + 2)/2 + 2 * (\lambda - 4)/4 = 7 \implies 2\lambda - 2 = 14 \implies \lambda = 8$

Therefore, the extreme point coordinates will be, $(x_1, y_1) = ((\lambda+2)/2, (\lambda-4)/4) = (5, 1)$

$f(5,1) = 5^2 + 4 * 1^2 - 2 * 5 + 8 * 1 = 25 + 4 - 10 + 8 = 27$

Taking 2 neighbouring point on the line $g(x,y) = x + 2y = 7$,

$f(3,2) = 3^2 + 4 * 2^2 - 2 * 3 + 8 * 2 = 9 + 16 - 6 + 16 = 35$

$f(1,3) = 1^2 + 4 * 3^2 - 2 * 1 + 8 * 3 = 2 + 36 - 2 + 24 = 60$

We can see the function f(x,y) has a minimum the point (5,1).

7. (1 point) The the minimum value of $f(x,y) = x^2 + 4y^2 - 2x + 8y$ subject to the constraint $x + 2y = 7$ occurs at the below point:

   A. (5,5)
   B. (-5,5)
   C. (1,5)
   D. (5,1)

Answer: D

Refer to the solution of the previous question

8. (1 point) The distance of the plane $x + y - 2z = 6$ from the origin is ____.

   Answer: 2.45, Range 2.30 to 2.60

   Given Squared Distance, $d^2 = f(x, y, z) = x^2 + y^2 + z^2$, $g(x, y, z) = x + y - 2z = 6$

   $\nabla f(x, y, z) = \begin{bmatrix} 2x \\ 2y \\ 2z \end{bmatrix}$, $\nabla g(x, y, z) = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$

   We find the values of x, y, z, λ that simultaneously satisfy the equations to get the extreme points $\nabla f(x, y, z) = \lambda \nabla g(x, y, z)$ and, $g(x, y, z) = x + y - 2z = 6$

   Solving, $\nabla f(x, y, z) = \lambda \nabla g(x, y, z)$

   $\implies \begin{bmatrix} 2x \\ 2y \\ 2z \end{bmatrix} = \lambda \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \implies x = \lambda/2, y = \lambda/2, z = -\lambda$

   Solving, $g(x, y, z) = x + y - 2z = 6 \implies \lambda/2 + \lambda/2 + 2\lambda = 6 \implies \lambda = 2$

   Therefore, the extreme point coordinates will be, $(x_1, y_1, z_1) = (\lambda/2, \lambda/2, -\lambda) = (1, 1, -2)$

   $f(1, 1, -2) = 1^2 + 1^2 + (-2)^2 = 1 + 1 + 4 = 6,$

   Taking 2 neighbouring point on the line $g(x, y, z) = x + y - 2z = 6,$

   $f(0, 0, -3) = 0^2 + 0^2 + (-3)^2 = 0 + 0 + 9 = 9$

   $f(2, 2, -1) = 2^2 + 2^2 + (-1)^2 = 4 + 4 + 1 = 9$

   We can see the function f(x,y,z) has a minimum at the point (1, 1, -2). The minimum distance of the plane $x + y - 2z = 6$ from the origin is, $d = \sqrt{6} = 2.45$

9. (1 point) The point on the plane $x + y - 2z = 6$ that is closest to the origin is

   A. $(0, 0, 0)$

   B. $(1, 1, 1)$

   C. $(-1, 1, 2)$

   D. $(1, 2, 0)$

   E. $(1, 1, -2)$

   Answer: E

   Refer to the solution of the previous question

10. (1 point) A box (cuboid shaped) is to made out of the cardboard with the total area of 24 cm^2. The maximum volume occupied by the box will be ____.

    Answer: 8, Range 7.50 to 8.50

    Given Volume, $V = f(x, y, z) = xyz,$

    Surface area, $g(x, y, z) = 2xy + 2yz + 2zx = 24 \implies g(x, y, z) = xy + yz + zx = 12$

$\nabla f(x,y,z) = \begin{bmatrix} yz \\ zx \\ xy \end{bmatrix}$, $\nabla g(x,y,z) = \begin{bmatrix} y+z \\ z+x \\ x+y \end{bmatrix}$

We find the values of x, y, z, λ that simultaneously satisfy the equations to get the extreme points $\nabla f(x,y,z) = \lambda \nabla g(x,y,z)$ and, $g(x,y,z) = xy + yz + zx = 12$

Solving, $\nabla f(x,y,z) = \lambda \nabla g(x,y,z)$

$\implies \begin{bmatrix} yz \\ zx \\ xy \end{bmatrix} = \lambda \begin{bmatrix} y+z \\ z+x \\ x+y \end{bmatrix} \implies xy + yz = yz + xz = xz + yz = xyz/\lambda \implies x = y = z = 0, 2\lambda$

Solving, $g(x,y,z) = xy + yz + zx = 12 \implies x = y = z = 2\lambda = \pm 2$

Since, x, y, z are length. This can not be a negative quantity. Therefore, the extreme point coordinates will be,

$(x_1, y_1, z_1) = (2\lambda, 2\lambda, 2\lambda) = (2,2,2)$, Volume, $V = f(2,2,2) = 2*2*2 = 8$

$(x_2, y_2, z_2) = (0,0,0)$, Volume, $V = f(0,0,0) = 0*0*0 = 0$

The box has maximum volume when it is a cube with edge $2cm$, and the maximum volume is $8cm^3$

11. (1 point) You are planning to setup a manufacturing business where the revenue (r) is a function of labour units (l), material units (m) and fixed cost (c), $r = l.m^2 + 2c$. You have an annual budget (b) of 1004 million rupees, $b = 2l + 16m + c$ to run the business. What would be maximum revenue that can be generated from the business in million rupees under the optimum combination of labour units (l), material units (m) and fixed cost(c).

    A. 1944

    B. 2036

    C. 2072

    D. 2080

Answer: A

Given

Revenue, $r = f(l,m,c) = l.m^2 + 2c$, Budget constraint, $g(l,m,c) = 2l + 16m + c = 1004$

$\nabla f(x,y,z) = \begin{bmatrix} m^2 \\ 2lm \\ 2 \end{bmatrix}$, $\nabla g(x,y,z) = \begin{bmatrix} 2 \\ 16 \\ 1 \end{bmatrix}$

We find the values of x, y, z, λ that simultaneously satisfy the equations to get the extreme points $\nabla f(l,m,c) = \lambda \nabla g(x,y,z)$ and, $g(l,m,c) = 2l + 16m + c = 1004$

Solving, $\nabla f(l,m,c) = \lambda \nabla g(x,y,z)$

$$\implies \begin{bmatrix} m^2 \\ 2lm \\ 2 \end{bmatrix} = \lambda \begin{bmatrix} 2 \\ 16 \\ 1 \end{bmatrix} \implies m = \pm 2, l = \pm 8, \lambda = 2$$

Solving, $g(l, m, c) = 2l + 16m + c = 1004 \implies c = 1004 - 2l - 16m$

Therefore, the extreme point coordinates will be,

$(l_1, m_1, c_1) = (8, 2, 956)$, Revenue, $r = f(8, 2, 956) = 8 * (2)^2 + 2 * 956 = 1944$

$(l_2, m_2, c_2) = (-8, -2, 1052)$, Revenue, $r = f(-8, -2, 1052) = -8 * (-2)^2 + 2 * 1052 = 2072$

Since, the configurations can not be negative. So, $l \geq 0, m \geq 0, c \geq 0$

We can see the maximum revenue is achieved under the configuration $l = 8, m = 2, c = 956$. The maximum revenue is 1944.

12. (1 point) The distance of the point on the sphere $x^2 + y^2 + z^2 = 3$ closest to the point (2,2,2) is ______.

Answer: 1.73, Range 1.50 to 2.00

Given Squared Distance, $d^2 = f(x, y, z) = (x-2)^2 + (y-2)^2 + (z-2)^2$, $g(x, y, z) = x^2 + y^2 + z^2 = 3$

$$\nabla f(x, y, z) = \begin{bmatrix} 2(x-2) \\ 2(y-2) \\ 2(z-2) \end{bmatrix}, \nabla g(x, y, z) = \begin{bmatrix} 2x \\ 2y \\ 2z \end{bmatrix}$$

We find the values of x, y, z, λ that simultaneously satisfy the equations to get the extreme points $\nabla f(x, y, z) = \lambda \nabla g(x, y, z)$ and, $g(x, y, z) = x^2 + y^2 + z^2 = 3$

Solving, $\nabla f(x, y, z) = \lambda \nabla g(x, y, z)$

$$\implies \begin{bmatrix} 2(x-2) \\ 2(y-2) \\ 2(z-2) \end{bmatrix} = \lambda \begin{bmatrix} 2x \\ 2y \\ 2z \end{bmatrix} \implies x = y = z = 2/(1-\lambda)$$

Solving, $g(x, y, z) = x^2 + y^2 + z^2 = 3 \implies 12/(1-\lambda)^2 = 3 \implies \lambda = 3, -1$

Therefore, the extreme point coordinates will be,

$(x_1, y_1, z_1) = (1, 1, 1)$, $d = \sqrt{(1-2)^2 + (1-2)^2 + (1-2)^2} = \sqrt{3} = 1.73$

$(x_2, y_2, z_2) = (-1, -1, -1)$, $d = \sqrt{((-1-2)^2 + (-1-2)^2 + (-1-2)^2} = 3\sqrt{3} = 5.20$

Taking 2 neighbouring point on the line $g(x, y, z) = x + y - 2z = 6$,

$(x_3, y_3, z_3) = (1, 1, -1)$, $d = \sqrt{(1-2)^2 + (1-2)^2 + (-1-2)^2} = \sqrt{11} = 3.32$

$(x_3, y_3, z_3) = (1, -1, -1)$, $d = \sqrt{(1-2)^2 + (-1-2)^2 + (-1-2)^2} = \sqrt{3} = 4.39$

We can see the point on the sphere closest to the point (2,2,2) is (1, 1, 1) and farthest from the point (2,2,2) is (-1, -1, -1)

13. (1 point) The distance of the point on the sphere $x^2 + y^2 + z^2 = 3$ farthest from the point (2,2,2) is ______.

Answer: 5.20, Range 5.00 to 5.50

Refer to the solution of the previous question

Course: Machine Learning - Foundations
Test Questions
Lecture Details: Week 9

---

1. (1 point) For a given point $x = 2$, is the function $f(x) = xe^{-3x}$ increasing or decreasing?

   A. increasing

   B. decreasing

   Answer: B

   If the first derivative of $f(x)$ at any point x is positive or negative, then we call $f(x)$ as increasing or decreasing respectively.
   $f'(x) = e^{-3x}(1 - 3x)$, at $x = 2$ we get $f'(2) = -0.012$. Hence decreasing.

2. (1 point) Find the value of a function$f(x) = x + 3x^2$ at its global minimum point.

   Answer: -0.0833

   range: -0.09,-0.08

   To get global minimum, $f'(x) = 0$. Therefore $1 + 6x = 0$, $x = -\frac{1}{6}$.

   $$f(\frac{1}{6}) = -0.0833$$

3. (1 point) Choose the largest interval of x in which a function $f(x) = xe^{x^2}$ is convex.

   A. (-0.5,0.5)

   B. (0, ∞)

   C. (∞,0)

   D. (1,0)

   Answer: B

   In order to find the interval over which $f(x)$ is convex, let us find where $f''(x) > 0$.

   $$f'(x) = 2x^2e^{x^2}$$
   $$f''(x) = e^{x^2}(4x^3 + 6x)$$

   Therefore $f''(x) > 0$ implies $e^{x^2}(4x^3 + 6x) > 0$.
   To satisfy this inequality, we want an interval of x for which $e^{x^2} > 0$ and $(4x^3 + 6x) > 0$.
   Because e raised to any power will be positive, first condition can be satisfied for any value of x.
   Second condition can be written as $2x(x^2 + 3) > 0$.
   $(x^2 + 3)$ will be greater than 0 for any values of x.
   $2x$ will be positive only when $x > 0$.
   Thus in interval notation, the largest interval of x for which $f(x)$ is convex is $(0, \inf)$.

4. (1 point) (Multiple select) Let the composition of two functions $f(x) = sin(x) - 2x^2 + 1$ and $g(x) = e^x$ be $h = fog$. At a point $x = 5$, Select the true statement(s).

   A. $h(x)$ is a convex function.

   B. $h(x)$ is a concave function.

   C. $h(x)$ is a non-decreasing function.

   D. $h(x)$ is a decreasing function.

   Answer: B,D

$$h(x) = f(g(x)) = sin(e^x) - 2e^{2x} + 1$$

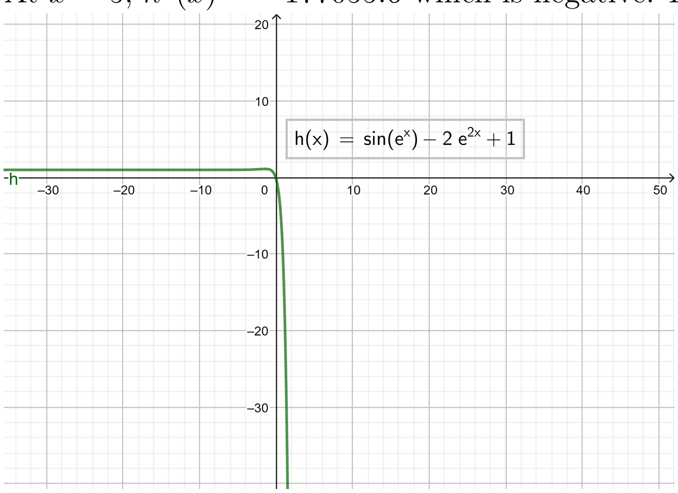
First derivative:

$$h'(x) = e^x cos(e^x) - 4e^{2x}$$

At $x = 5$; $h'(x) = -88232.28$ which is less than 0. Hence h(x) is decreasing.
Second derivative:

$$h''(x) = e^{2x}(cos(e^x) - sin(e^x)) - 8e^{2x}$$

At $x = 5$; $h''(x) = -177055.6$ which is negative. Hence h(x) is concave.



5. (2 points) Find the minimum value of $f(x, y) = x + y$ subject to $x^2 + y^2 = 1$, where x, y are the coordinates of points on the circumference of the unit circle.

   Answer: -1.414

range: -1,-2 The Lagrangian function for the problem is

$$L(x, y, \lambda) = x + y + \lambda x^2 + \lambda y^2 - \lambda$$

Take partial derivatives with respect to x,y, λ and equate to zero.

$$\frac{\partial L}{\partial x} = 2\lambda x + 1 = 0$$

$$x = -\frac{1}{2\lambda} \tag{1}$$

$$\frac{\partial L}{\partial y} = 2\lambda y + 1 = 0$$

$$y = -\frac{1}{2\lambda} \tag{2}$$

$$\frac{\partial L}{\partial \lambda} = x^2 + y^2 = 1 \tag{3}$$

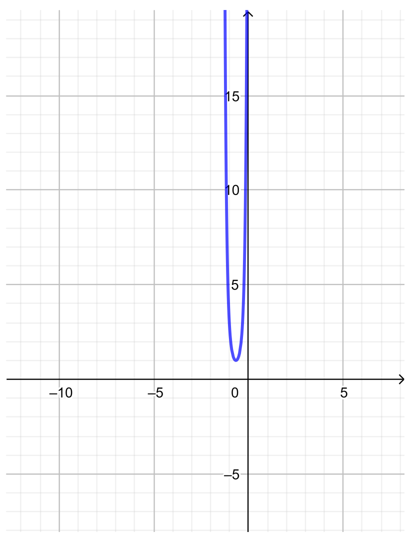
Substituting (1) and (2) in (3), we get

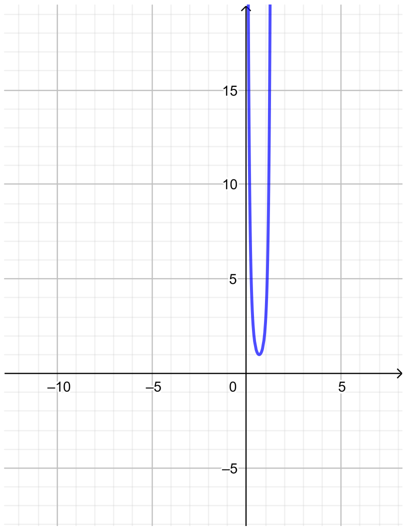
$$\frac{1}{4\lambda^2} + \frac{1}{4\lambda^2} = 1$$

$$\lambda = \pm\frac{1}{\sqrt{2}} \tag{4}$$

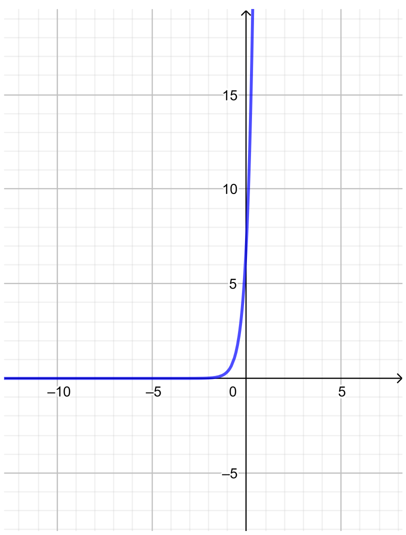
Using (3) in (1) and (2) we get, $x = \mp\frac{1}{\sqrt{2}}$ and $y = \mp\frac{1}{\sqrt{2}}$. Since we want to minimize $f(x,y)$, we shall consider $x = -\frac{1}{\sqrt{2}}$ and $y = -\frac{1}{\sqrt{2}}$.
Minimum value of $f(x,y) = -1.414$.

6. (1 point) For the functions $g(x) = (3x+2)^2$ and $f(x) = e^x$, select the plot that corresponds to the correct composition $h = fog$.

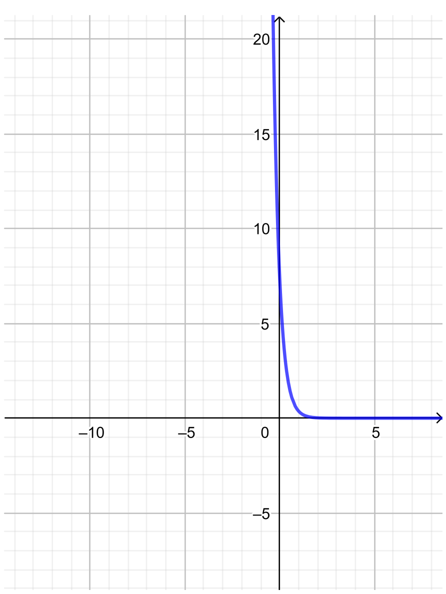
A. 



B.



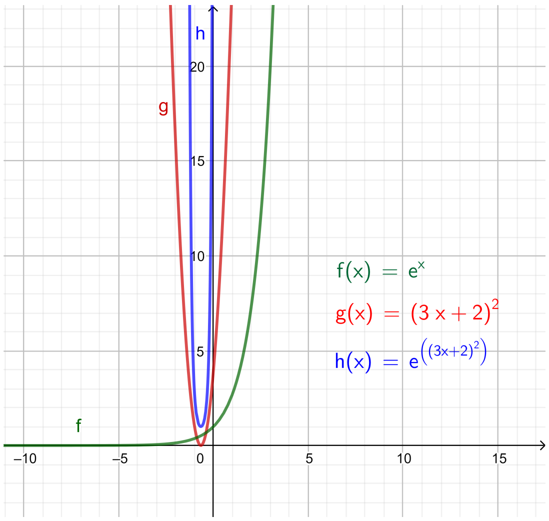
C.



D.

Answer: A

The functions f(x), g(x) and their composition h(x) are plotted as follows:



(Common data for Q7, Q8) Linear programming deals with the problem of finding a vector x that minimizes a given linear function $c^T x$, where x ranges over all vectors ($x \geq 0$) satisfying a given system of linear equations $Ax = b$. Here A is a $m \times n$ matrix, $c, x \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$.

7. (1 point) Choose the correct dual program with y as the dual variable for the above linear program from the following.

    A.
    $$\min_x by \quad \text{subject to} \quad A^T y \geq c$$

    B.
    $$\max_x b^T y \quad \text{subject to} \quad A^T y \leq c$$

    C.
    $$\max_x b^T y \quad \text{subject to} \quad A^T y \geq c$$

    D.
    $$\max_x by \quad \text{subject to} \quad A^T y \leq c$$

   Answer: B

8. (1 point) From the below given statements regarding constraints and decision variables related to the primal and dual problems of the linear program, choose the correct statement.

    A. Primal problem has m constraints and m decision variables whereas dual problem has n constraints and n decision variables.

    B. Primal problem has m constraints and n decision variables whereas dual problem has n constraints and m decision variables.

C. Primal problem has n constraints and n decision variables whereas dual problem has m constraints and m decision variables.

D. Primal problem has n constraints and m decision variables whereas dual problem has m constraints and n decision variables.

Answer: B

9. (2 points) Let a set of data points with five samples and two features per sample be $X = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 4 & 2.5 \\ 6 & 4 \\ 7.5 & 5 \end{bmatrix}$ and the corresponding labels be $y = \begin{bmatrix} 1.5 \\ 2 \\ 2.5 \\ 3 \\ 4 \end{bmatrix}$. Perform linear regression on this data set and choose the optimal solution for w^* to minimize the sum of squares error.

A. $\begin{bmatrix} 0.2763 \\ 1.2039 \end{bmatrix}$

B. $\begin{bmatrix} 0.0691 \\ 0.3010 \end{bmatrix}$

C. $\begin{bmatrix} 0.1382 \\ 0.6019 \end{bmatrix}$

D. $\begin{bmatrix} 0.0276 \\ 0.1204 \end{bmatrix}$

Answer: C

optimal $w^* = (X^T X)^{-1}(X^T y)$

$X^T X = \begin{bmatrix} 113.25 & 79.5 \\ 79.5 & 60.25 \end{bmatrix}$

$X^T y = \begin{bmatrix} 63.5 \\ 47.25 \end{bmatrix}$

$w^* = \begin{bmatrix} 0.1382 \\ 0.6019 \end{bmatrix}$

A, B and D are scalar multiples of C.

(Common data for Q10, Q11) Krishna runs a steel fabrication industry and produces steel products. He regularly purchases raw steel for Rs.500 per ton. His revenue is modeled by a function $R(s) = 100\sqrt{s}$, where s is the tons of steel purchased. His budget for steel purchase is Rs.150000.

10. (1 point) Using Lagrangian function, find the amount of raw steel to be purchased to get maximum revenue?

Answer: 300

range: 298, 302

Lagrangian function: $L(s, \lambda) = 100s^{0.5} + \lambda(s - 150000)$

$$L(s, \lambda) = 100s^{0.5} + s\lambda - \lambda 150000$$

Take partial derivatives with respect to s and λ and equate to zero.

$$\frac{\partial L}{\partial s} = 0 = 50s^{-0.5} + \lambda 500 = 0$$

$$\frac{\partial L}{\partial \lambda} = 0 = s500 - 150000 = 0$$

$$s = 300$$

11. (1 point) What is the maximum revenue value in Rs?

Answer: 1732.05

range: 1730,1734

Maximum revenue: $100 * \sqrt{300} = 1732.05$

12. (1 point) Consider a vector $\hat{w} = \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix}$ in $\mathbb{R}^3$. In $\mathbb{R}^3$, there are many unit vectors. Use Lagrange method to find the unit vector which gives the minimum dot product.

A. $\hat{u} = \frac{1}{2\lambda} \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix}$, with $\lambda \geq 0$

B. $\hat{u} = \frac{-1}{3\lambda} \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix}$, with $\lambda \geq 0$

C. $\hat{u} = \frac{-1}{4\lambda} \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix}$, with $\lambda \geq 0$

D. $\hat{u} = \frac{-1}{2\lambda} \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix}$, with $\lambda \geq 0$

Answer: D

Let unit vector be $\vec{u} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$.

Objective is to minimize $\vec{u}.\vec{w}$ subject to $x^2 + y^2 + z^2 = 1$.

$$f(x, y, z) = \vec{u}.\vec{w} = 2x + 4y + 3z$$

Lagrangian function can be written as follows:

$$L(x, y, z, \lambda) = (2x + 4y + 3z) + \lambda(x^2 + y^2 + z^2 - 1)$$

Take partial derivatives with respect to x, y, z and λ and equate to zero.

$$\frac{\partial L}{\partial x} = 2 + 2\lambda x = 0$$

$$x = \frac{-1}{2\lambda}2$$

$$\frac{\partial L}{\partial y} = 4 + 2\lambda y = 0$$

$$y = \frac{-1}{2\lambda}4$$

$$\frac{\partial L}{\partial z} = 3 + 2\lambda z = 0$$

$$z = \frac{-1}{2\lambda}3$$

(Common data for Q13, Q14) Solve the following linear program using KKT conditions.

$$\text{minimize} \quad v = 24y_1 + 60y_2$$

subject to

$$0.5y_1 + y_2 \geq 6$$

$$2y_1 + 2y_2 \geq 14$$

$$y_1 + 4y_2 \geq 13$$

$$y_1 \geq 0, y_2 \geq 0$$

13. (2 points) Choose the optimal solution for $[y_1^*, y_2^*]$
    - A. [8,2.5]
    - B. [10,6]
    - C. [11,0.5]
    - D. [10.5,1]

Answer: C

$[y_1^*, y_2^*] = [11, 0.5]$, satisfies all the constraints and gives minimum value of v.

14. (1 point) Enter the minimum value of v.

Answer: 294

$$v = 24(11) + 60(0.5) = 294$$

Course: Machine Learning - Foundations
Week 10 Questions

TEST QUESTIONS

1. (1 point) **(Multiple Select)** For three events, A, B, and C, with $P(C) > 0$, Which of the following is/are correct?

   A. $P(A^c|C)$= 1 - $P(A|C)$

   B. $P(\phi|C) = 0$

   C. $P(A|C) \leq 1$

   D. if $A \subset B$ then $P(A|C) \leq P(B|C)$

   Answer: A, B, C, D

2. (2 points) **(Multiple Select)** Let the random experiment be tossing an unbiased coin two times. Let A be the event that the first toss results in a head, B be the event that the second toss results in a tail and C be the event that on both the tosses, the coin landed on the same side. Choose the correct statements from the following:

   A. A and C are independent events.

   B. A and B are independent events.

   C. B and C are independent events.

   D. A, B, and C are independent events.

   Answer: A, B, C

   Solution:

   $A = \{HT, HH\}$

   $B = \{HT, TT\}$

   $C = \{TT, HH\}$

   $P(A) = \frac{1}{2}$

   $P(B) = \frac{1}{2}$

   $P(C) = \frac{1}{2}$

   $P(A \cap B) = \{HT\}$

   $P(C \cap B) = \{TT\}$

   $P(A \cap C) = \{HH\}$

$P(A \cap B) = P(A) \times P(B)$ Hence, option B is correct

$P(A \cap C) = P(A) \times P(C)$ Hence, option A is correct

$P(C \cap B) = P(C) \times P(B)$ Hence, option C is correct

3. (2 points) **(Multiple Select)** If $A_1, A_2, A_3, , A_n$ are non empty disjoint sets and subsets of sample space S, and a set A_{n+1} is also a subset of S, then which of the following statements are true?

   A. The sets $A_1 \cap A_{n+1}, A_2 \cap A_{n+1}, A_3 \cap A_{n+1}, , A_n \cap A_{n+1}$ are disjoint.

   B. If A_{n+1}, A_n are disjoint then $A_1, A_2, , A_{n-1}$ are disjoint with A_{n+1}.

   C. The sets $A_1, A_2, A_3, , A_n, \phi$ are disjoint.

   D. The sets $A_1, A_2, A_3, , A_n, S$ are disjoint.

   Answer: A, C

4. (3 points) A triangular spinner having three outcomes can lands on one of the numbers 0, 1 and 2 with probabilities shown in table.

Outcome	0	1	2
Probability	0.7	0.2	0.1

Table 1: Table 10.2: Probability distribution

The spinner is spun twice. The total of the numbers on which it lands is denoted by X. The the probability distribution of X is.

A.

x	2	3	4	5	6
$P(X = x)$	$\frac{49}{100}$	$\frac{28}{100}$	$\frac{1}{100}$	$\frac{4}{100}$	$\frac{18}{100}$

B.

x	2	3	4	5	6
$P(X = x)$	$\frac{28}{100}$	$\frac{49}{100}$	$\frac{18}{100}$	$\frac{1}{100}$	$\frac{4}{100}$

C.

x	0	1	2	3	4
$P(X = x)$	$\frac{49}{100}$	$\frac{28}{100}$	$\frac{18}{100}$	$\frac{4}{100}$	$\frac{1}{100}$

D.

x	2	3	4	5	6
$P(X = x)$	$\frac{28}{100}$	$\frac{49}{100}$	$\frac{18}{100}$	$\frac{4}{100}$	$\frac{1}{100}$

Answer: C

5. (1 point) When throwing a fair die, what is the variance of the number of throws needed to get a 1?

Answer: 30

Solution:

$$= Var(X) = \frac{1-p}{p^2}$$

$$= \frac{1-\frac{1}{6}}{\frac{1}{6}^2}$$

$$= 30$$

6. (1 point) Joint pmf of two random variables X and Y are given in Table

$x \backslash y$	1	2	3	$f_X(x)$
1	0.05	0	a_1	0.15
2	0.1	0.2	a_3	a_2
3	a_4	0.2	a_5	0.45
$f_Y(y)$	0.3	0.4	a_6	

Find the value of $f_{Y|X=3}(1)$ i.e ($P(Y=1|X=3)$)

Answer: 0.22

Solution:

$\sum f_{XY}(x,y) = 1$ .... (i)

$f_X(x) = \sum\limits_{y \in R_y} f_{XY}(x,y)$ ....(ii)

$f_Y(y) = \sum\limits_{x \in R_X} f_{XY}(x,y)$ ....(iii)

Hence, $a_1 = 0.10$, $a_2 = 0.40$, $a_3 = 0.1$, $a_4 = 0.15$, $a_5 = 0.1$, $a_6 = 0.3$

$$f_{Y|X=3}(1) = \frac{f_{XY}(1,3)}{f_X(3)} = \frac{0.1}{0.45} = 0.22$$

7. (1 point) **(Multiple Select)** Which of the following options is/are correct?
   A. If $Cov[X, Y] = 0$, then X and Y are independent random variables.
   B. $Cov[X, X] = Var(X)$
   C. If X and Y are two independent random variables and $Z = X + Y$ then $f_Z(z) = \sum_x f_X(x) \times f_Y(z - x)$
   D. If X and Y are two independent random variables and $Z = X + Y$ then $f_Z(z) = \sum_y f_X(x) \times f_Y(z - x)$

Answer: B, C

Solution:

Option B

$Cov[X, X]$ is the covariance between X and X i.e $Var(X)$

Option C is correct from its definition.

8. (1 point) **(Multiple Select)**A discrete random variables X has the cumulative distribution function is defined as follows.

$$F_X(x) = \begin{cases} \dfrac{x^3 + k}{40}, & \text{for } x = 1, 2, 3 \end{cases}$$

Which of the following options is/are correct for $F(x)$ as given?
   A. $k = 17$
   B. $Var(X) = \dfrac{259}{320}$
   C. $k = 13$
   D. $Var(X) = \dfrac{249}{310}$

Answer: B, C

Solution:

For k

$F_X(3) = 1$

$\dfrac{x^3 + k}{40} = 1$

Solving above equation to get $k = 13$

To calculate the variance, first calculate the probability distribution of X

We will get

$P(X = 1) = \dfrac{14}{40}$

$P(X = 2) = \dfrac{7}{40}$

$P(X = 3) = \dfrac{19}{40}$

Now easily with $Var(X)$ equation we will get $Var(X) = \dfrac{259}{320}$

9. (1 point) In a game of Ludo, Player A needs to repeatedly throw an unbiased die till he gets a 6. What is the probability that he needs fewer than 4 throws? (Answer the question correct to two decimal points.)

Solution:

$P(6) = \dfrac{1}{6}$

As it resembles geometric distribution. Hence,
$\sum_{n=1}^{3} \dfrac{1}{6} \times (1 - \dfrac{1}{6})^5 = 0.6$

10. (1 point) **(Multiple Select)** Let X and Y be two random variables with joint PMF $f_{XY}(x, y)$ given in Table 10.3.

x \ y	0	1	2
0	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{8}$
1	$\frac{1}{8}$	$\frac{1}{6}$	$\frac{1}{6}$

Table 10.3: Joint PMF of X and Y.

Which of the following options is/are correct for $f_{XY}(x, y)$ given in Table 10.1.

A. $P(X=0, Y \leq 1) = \dfrac{5}{12}$

B. $P(X=0, Y \leq 1) = \dfrac{7}{12}$

C. X and Y are independent.

D. X and Y are dependent.

Answer: A, D

11. (1 point) A discrete random variables X has the probability function as given in table 10.4.

x	1	2	3	4	5	6
$P(X)$	a	a	a	b	b	0.3

Table 2: Table 10.4: Probability distribution

If $E(X) = 4.2$, then evaluate $a + b$

Answer: 0.3

$\sum P(X = x) = 1$

$3a + 2b = 0.7$

$E(X) = \sum P(X = x_i) \times x_i$

$6a + 9b = 2.4$

Solving both equations, we get $a = 0.1$ and $b = 0.2$

12. (1 point) A discrete random variable X has the probability function as follows.

$$P(X = x) = \begin{cases} k \times (1-x)^2, & \text{for } x = 1, 2, 3 \\ 0, & \text{otherwise} \end{cases}$$

Evaluate $E(X)$

Answer: 2.8

Solution:

$\sum P(X = x) = 1$

$k + 4k = 1$

$k = 0.2$

$E(X) = \sum P(X = x_i) \times x_i$

$0.2 \times 2 + 0.8 \times 3$

$0.4 + 2.4 = 2.8$

Course: Machine Learning - Foundations
Week 12 Questions

GRADED QUESTIONS

1. (1 point) Let X be random variable of binomial(n, p). Using Markov's inequality, find an upper bound on $P(X \geq \alpha n)$, where $p < \alpha < 1$. Evaluate the upper bound for $p = \dfrac{1}{3}$ and $\alpha = \dfrac{3}{4}$.

   A. $\dfrac{2}{3}$

   B. $\dfrac{4}{9}$

   C. $\dfrac{3}{5}$

   D. $\dfrac{4}{11}$

   Answer: B

   Solution: Note that X is a non-negative random variable and $E(X) = np$. Applying Markov's inequality, we obtain,

   $P(X \geq \alpha n) \leq \dfrac{E(X)}{\alpha n} = \dfrac{pn}{\alpha n} = \dfrac{p}{\alpha}$.

   for $p = \dfrac{1}{3}, \alpha = \dfrac{3}{4}$ we obtain,

   $P(X \geq \dfrac{3n}{4}) \leq \dfrac{4}{9}$.

2. (1 point) Let X be random variable of binomial(n, p). Using Chebyshev's inequality, find an upper bound on $P(X \geq \alpha n)$, where $p < \alpha < 1$. Evaluate upper the bound for $p = \dfrac{1}{3}$ and $\alpha = \dfrac{3}{4}$ and $n = 4$

   Answer: 0.32

   Solution:

   One way to obtain a bound is to write

   $P(X \geq \alpha n) = P(X - np \geq \alpha n - np)$
   $\leq (|X - np| \geq n\alpha - np)$
   $\leq \dfrac{Var(X)}{(n\alpha - np)^2}$

Putting the values of α, p and n

we will get upper bound as $\frac{288}{225n}$

3. (1 point) Suppose X is a non-negative random variable with expectation 60 and Standard deviation 5. What can we say about the best upper bound of $P(X \geq 70)$ (Hint: Use Chebyshev's inequality)?

Answer: 0.25

Since X is non-negative, we could just apply Chebyshev's inequality,

$P(X \geq 70) = P(X - 60 \geq 10) \leq \frac{5^2}{10^2} = \frac{1}{4} = 0.25$

If Y follows $\mathcal{N}(\mu, \sum)$, where Y is a vector that is, $Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$ and $\mu = \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix}$ and $\sum = \begin{pmatrix} 6 & 1 & -2 \\ 1 & 13 & 4 \\ -2 & 4 & 4 \end{pmatrix}$

From the above information answer questions

4. (points) Suppose $Z = CY$, where $C = (2, -1, 3)$. Find the distribution of $Z = 2y_1 - y_2 + 3y_3$
   A. $\mathcal{N}(17, 21)$
   B. $\mathcal{N}(17, 15)$
   C. $\mathcal{N}(17, 17)$
   D. $\mathcal{N}(15, 21)$

Answer: A

Solution:
$E(Z) = E(CY) = CE(Y)$
$Var(Z) = Var(CY) = CVar(Y)\bar{C}$

5. (points) If $Z = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = DY$, Where $D = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \end{pmatrix}$ Then Find the distribution of Z

Answer: N($\begin{pmatrix} 8 \\ 10 \end{pmatrix}$, $\begin{pmatrix} 29 & -1 \\ -1 & 9 \end{pmatrix}$)

Solution:

$E(Z) = E(DY) = DE(Y)$
$Var(Z) = Var(DY) = DVar(Y)\bar{D}$

6. (points) Find the maximum likelihood estimate for the parameter λ of a poisson distribution of sample values $x_1, x_2, ...x_n$. Here $\bar{x}$ represent the mean value of the sample values $x_1, x_2, ...x_n$?

   A. $\lambda = \bar{x}$

   B. $\lambda = 2\bar{x}$

   C. $\lambda = 3\bar{x}$

   D. $\lambda = 4\bar{x}$

Answer: A

Solution:

The probability function of the poisson distribution with parameter given by:

$P(X = x) = f(x, \lambda) = \dfrac{e^{-\lambda}\lambda^x}{x!}; x = 0, 1, 2, ...$

$L = \prod_{i=1}^{n} f(x_i, \lambda) = \dfrac{e^{-n\lambda}\lambda^{\sum_i^n x_i}}{x_1!x_2!.....x_{n!}}$

$logL = -n\lambda + n\bar{x}log\lambda - \sum_i^n log(x_i!)$

The likelihood equation for estimating λ is:

$\dfrac{\partial logL}{\partial \lambda} = 0$

$-n + \dfrac{n\bar{x}}{\lambda} = 0$

$\lambda = \bar{x}$

7. (points) If X is the number scored in a throw of a fair die. Then which of the following options are correct. (Hint: Use Chebychev's inequality to solve).

   A. $P(|X - \mu| > 2.5) < 0.47$

B. $P(|X-\mu| > 2.5) > 0.47$

C. $P(|X-\mu| \leq 2.5) \leq 0.47$

D. $P(|X-\mu| \leq 2.5) < 0.47$

Answer: A

Solution:
Here X is a random variable which takes the values 1, 2, ...6 each with probability $\frac{1}{6}$.
Hence $E(X) = 3.5$

$E(X^2) = \frac{91}{6}$

$Var(X) = \frac{35}{12}$

For $k > 0$, Chebychev's inequality gives $P(|X-\mu| > k) \leq \frac{Var}{k^2}$

Choosing $k = 2.5$, we get $P(|X-\mu| > k) \leq 0.47$

8. (points) In random sampling from normal distribution $\mathcal{N}(\mu, \sigma^2)$, find the maximum likelihood estimators for σ^2 when μ is known

A. $\sigma = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n}$

B. $\sigma^2 = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n}$

C. $\sigma^2 = \frac{\sum_{i=1}^{n}(x_i-\mu)}{n}$

D. $\sigma^2 = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2n}$

Answer: B

Solution:
$L = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp -\frac{1}{2\sigma^2}(x_i-\mu)^2 = (\frac{1}{\sigma\sqrt{2\pi}})^n \exp -\sum_{i=1}^{n} -\frac{1}{2\sigma^2}(x_i-\mu)^2$

Taking log on both sides,

$logL = -\frac{n}{2}log(2\pi) - \frac{n}{2}log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$

When μ is known, the likelihood equation for estimating σ^2 is

$\frac{\partial logL}{\partial \sigma^2} = 0$

Taking partial differentiation an solving we will get option B as correct answer.

9. (points) The Central Limit Theorem says that the sampling distribution of the sample mean is approximately normal if

   A. all selected samples $x_1, x_2, x_3, ...$ are independent.

   B. the sample size is large.

   C. Both A and B

   D. Always

Answer: B

Solution:
From definition of central limit theorem, option B is correct.

Course: Machine Learning - Foundations
Week 10 Questions

TEST QUESTIONS

1. (1 point) **(Multiple Select)** For three events, A, B, and C, with $P(C) > 0$, Which of the following is/are correct?

   A. $P(A^c|C)$= 1 - $P(A|C)$

   B. $P(\phi|C) = 0$

   C. $P(A|C) \leq 1$

   D. if $A \subset B$ then $P(A|C) \leq P(B|C)$

   Answer: A, B, C, D

2. (2 points) **(Multiple Select)** Let the random experiment be tossing an unbiased coin two times. Let A be the event that the first toss results in a head, B be the event that the second toss results in a tail and C be the event that on both the tosses, the coin landed on the same side. Choose the correct statements from the following:

   A. A and C are independent events.

   B. A and B are independent events.

   C. B and C are independent events.

   D. A, B, and C are independent events.

   Answer: A, B, C

   Solution:

   $A = \{HT, HH\}$

   $B = \{HT, TT\}$

   $C = \{TT, HH\}$

   $P(A) = \dfrac{1}{2}$

   $P(B) = \dfrac{1}{2}$

   $P(C) = \dfrac{1}{2}$

   $P(A \cap B) = \{HT\}$

   $P(C \cap B) = \{TT\}$

   $P(A \cap C) = \{HH\}$

$P(A \cap B) = P(A) \times P(B)$ Hence, option B is correct

$P(A \cap C) = P(A) \times P(C)$ Hence, option A is correct

$P(C \cap B) = P(C) \times P(B)$ Hence, option C is correct

3. (2 points) **(Multiple Select)** If $A_1, A_2, A_3, , A_n$ are non empty disjoint sets and subsets of sample space S, and a set A_{n+1} is also a subset of S, then which of the following statements are true?

   A. The sets $A_1 \cap A_{n+1}, A_2 \cap A_{n+1}, A_3 \cap A_{n+1}, , A_n \cap A_{n+1}$ are disjoint.

   B. If A_{n+1}, A_n are disjoint then $A_1, A_2, , A_{n-1}$ are disjoint with A_{n+1}.

   C. The sets $A_1, A_2, A_3, , A_n, \phi$ are disjoint.

   D. The sets $A_1, A_2, A_3, , A_n, S$ are disjoint.

   Answer: A, C

4. (3 points) A triangular spinner having three outcomes can lands on one of the numbers 0, 1 and 2 with probabilities shown in table.

Outcome	0	1	2
Probability	0.7	0.2	0.1

Table 1: Table 10.2: Probability distribution

The spinner is spun twice. The total of the numbers on which it lands is denoted by X. The the probability distribution of X is.

A.

x	2	3	4	5	6
$P(X = x)$	$\frac{49}{100}$	$\frac{28}{100}$	$\frac{1}{100}$	$\frac{4}{100}$	$\frac{18}{100}$

B.

x	2	3	4	5	6
$P(X = x)$	$\frac{28}{100}$	$\frac{49}{100}$	$\frac{18}{100}$	$\frac{1}{100}$	$\frac{4}{100}$

C.

x	0	1	2	3	4
$P(X = x)$	$\frac{49}{100}$	$\frac{28}{100}$	$\frac{18}{100}$	$\frac{4}{100}$	$\frac{1}{100}$

D.

x	2	3	4	5	6
$P(X = x)$	$\frac{28}{100}$	$\frac{49}{100}$	$\frac{18}{100}$	$\frac{4}{100}$	$\frac{1}{100}$

Answer: C

5. (1 point) When throwing a fair die, what is the variance of the number of throws needed to get a 1?

Answer: 30

Solution:

$$= Var(X) = \frac{1-p}{p^2}$$

$$= \frac{1-\frac{1}{6}}{\frac{1}{6}^2}$$

$$= 30$$

6. (1 point) Joint pmf of two random variables X and Y are given in Table

$x \backslash y$	1	2	3	$f_X(x)$
1	0.05	0	a_1	0.15
2	0.1	0.2	a_3	a_2
3	a_4	0.2	a_5	0.45
$f_Y(y)$	0.3	0.4	a_6	

Find the value of $f_{Y|X=3}(1)$ i.e $(P(Y = 1|X = 3))$

Answer: 0.22

Solution:

$\sum f_{XY}(x, y) = 1$ .... (i)

$f_X(x) = \sum\limits_{y \in R_y} f_{XY}(x, y)$ ....(ii)

$f_Y(y) = \sum\limits_{x \in R_X} f_{XY}(x, y)$ ....(iii)

Hence, $a_1 = 0.10$, $a_2 = 0.40$, $a_3 = 0.1$, $a_4 = 0.15$, $a_5 = 0.1$, $a_6 = 0.3$

$$f_{Y|X=3}(1) = \frac{f_{XY}(1,3)}{f_X(3)} = \frac{0.1}{0.45} = 0.22$$

7. (1 point) **(Multiple Select)** Which of the following options is/are correct?

   A. If $Cov[X, Y] = 0$, then X and Y are independent random variables.

   B. $Cov[X, X] = Var(X)$

   C. If X and Y are two independent random variables and $Z = X + Y$ then $f_Z(z) = \sum_x f_X(x) \times f_Y(z - x)$

   D. If X and Y are two independent random variables and $Z = X + Y$ then $f_Z(z) = \sum_y f_X(x) \times f_Y(z - x)$

Answer: B, C

Solution:

Option B

$Cov[X, X]$ is the covariance between X and X i.e $Var(X)$

Option C is correct from its definition.

8. (1 point) **(Multiple Select)**A discrete random variables X has the cumulative distribution function is defined as follows.

$$F_X(x) = \begin{cases} \dfrac{x^3 + k}{40}, & \text{for } x = 1, 2, 3 \end{cases}$$

Which of the following options is/are correct for $F(x)$ as given?

   A. $k = 17$

   B. $Var(X) = \dfrac{259}{320}$

   C. $k = 13$

   D. $Var(X) = \dfrac{249}{310}$

Answer: B, C

Solution:

For k

$F_X(3) = 1$

$\frac{x^3 + k}{40} = 1$

Solving above equation to get $k = 13$

To calculate the variance, first calculate the probability distribution of X

We will get

$P(X = 1) = \frac{14}{40}$

$P(X = 2) = \frac{7}{40}$

$P(X = 3) = \frac{19}{40}$

Now easily with $Var(X)$ equation we will get $Var(X) = \frac{259}{320}$

9. (1 point) In a game of Ludo, Player A needs to repeatedly throw an unbiased die till he gets a 6. What is the probability that he needs fewer than 4 throws? (Answer the question correct to two decimal points.)

   Solution:

   $P(6) = \frac{1}{6}$

   As it resembles geometric distribution. Hence,
   $\sum_{n=1}^{3} \frac{1}{6} \times (1 - \frac{1}{6})^5 = 0.6$

10. (1 point) **(Multiple Select)** Let X and Y be two random variables with joint PMF $f_{XY}(x, y)$ given in Table 10.3.

x \ y	0	1	2
0	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{8}$
1	$\frac{1}{8}$	$\frac{1}{6}$	$\frac{1}{6}$

Table 10.3: Joint PMF of X and Y.

Which of the following options is/are correct for $f_{XY}(x, y)$ given in Table 10.1.

A. $P(X=0, Y \leq 1) = \frac{5}{12}$

B. $P(X=0, Y \leq 1) = \frac{7}{12}$

C. X and Y are independent.

D. X and Y are dependent.

Answer: A, D

11. (1 point) A discrete random variables X has the probability function as given in table 10.4.

x	1	2	3	4	5	6
$P(X)$	a	a	a	b	b	0.3

Table 2: Table 10.4: Probability distribution

If $E(X) = 4.2$, then evaluate $a + b$

Answer: 0.3

$\sum P(X = x) = 1$

$3a + 2b = 0.7$

$E(X) = \sum P(X = x_i) \times x_i$

$6a + 9b = 2.4$

Solving both equations, we get $a = 0.1$ and $b = 0.2$

12. (1 point) A discrete random variable X has the probability function as follows.

$$P(X = x) = \begin{cases} k \times (1-x)^2, & \text{for } x = 1, 2, 3 \\ 0, & \text{otherwise} \end{cases}$$

Evaluate $E(X)$

Answer: 2.8

Solution:

$\sum P(X = x) = 1$

$k + 4k = 1$

$k = 0.2$

$E(X) = \sum P(X = x_i) \times x_i$

$0.2 \times 2 + 0.8 \times 3$

$0.4 + 2.4 = 2.8$