IAE ACADEMY INSTITUTE FOR ENGINEERS

GATE DA · 2025

IAE Academy free study material · Updated 19 October 2026

GATE DA 2025 Solved Paper

INSIDE

- All 65 questions, in the order of the paper
- The official answer to every question
- A step-by-step solution to every question
- The answer key on one page

Question paper and official answer key: IIT Roorkee

Online, with the questions linked to their topic notes: **iaeacademy.in/study/papers/da-2025**

FREE COUNSELLING

Preparing for GATE DA? Talk to IAE Academy.

Or call **+91 84204 36524**

Every day, 10:30 am to 7:30 pm

Chat on WhatsApp

Free. Not for sale. Pass it on.

About this paper

65 questions, 100 marks. Question types: 35 MCQ · 18 MSQ · 12 NAT.

Official answer key: IIT Roorkee · https://gate2025.iitr.ac.in/doc/2025/2025_Key/DA_Keys.pdf

We placed each question under a subject and topic ourselves; the exam paper does not label them.

This paper online, with the questions linked to their topic notes:
iaeacademy.in/study/papers/da-2025

Answer key

Q	TYPE	MARKS	ANSWER
Q.1	MCQ	1	**A**
Q.2	MCQ	1	**B**
Q.3	MCQ	1	**C**
Q.4	MCQ	1	**A**
Q.5	MCQ	1	**C**
Q.6	MCQ	2	**B**
Q.7	MCQ	2	**B**
Q.8	MCQ	2	**A**
Q.9	MCQ	2	**C**
Q.10	MCQ	2	**A**
Q.11	MCQ	1	**C**
Q.12	MCQ	1	**B**
Q.13	MCQ	1	**C**
Q.14	MCQ	1	**A**
Q.15	MCQ	1	**A**
Q.16	MCQ	1	**C**
Q.17	MCQ	1	**C**
Q.18	MCQ	1	**D**
Q.19	MCQ	1	**A**
Q.20	MCQ	1	**A**
Q.21	MCQ	1	**A**
Q.22	MCQ	1	**A**
Q.23	MCQ	1	**A**
Q.24	MSQ	1	**A, B, C**
Q.25	MSQ	1	**B, D**
Q.26	MSQ	1	**C, D**
Q.27	MSQ	1	**C**
Q.28	MSQ	1	**A, B**
Q.29	MSQ	1	**A, C**
Q.30	MSQ	1	**B, D**
Q.31	NAT	1	**0.25**
Q.32	NAT	1	**0.5**
Q.33	NAT	1	**3**
Q.34	NAT	1	**0.285 to 0.287**
Q.35	NAT	1	**0.39 to 0.41**
Q.36	MCQ	2	**B**
Q.37	MCQ	2	**D**
Q.38	MCQ	2	**A**
Q.39	MCQ	2	**C**
Q.40	MCQ	2	**A**
Q.41	MCQ	2	**A**
Q.42	MCQ	2	**A**
Q.43	MCQ	2	**C**
Q.44	MCQ	2	**C**
Q.45	MCQ	2	**C**
Q.46	MCQ	2	**D**
Q.47	MCQ	2	**B**
Q.48	MSQ	2	**A, C**
Q.49	MSQ	2	**C, D**
Q.50	MSQ	2	**A, B**
Q.51	MSQ	2	**B, C, D**
Q.52	MSQ	2	**A, D**
Q.53	MSQ	2	**B, C**
Q.54	MSQ	2	**A, C**
Q.55	MSQ	2	**B, C**
Q.56	MSQ	2	**B, C**
Q.57	MSQ	2	**B, D**
Q.58	MSQ	2	**A, C, D**
Q.59	NAT	2	**0**
Q.60	NAT	2	**100**
Q.61	NAT	2	**66.6 to 66.7**
Q.62	NAT	2	**1**
Q.63	NAT	2	**160**
Q.64	NAT	2	**24**
Q.65	NAT	2	**75**

General Aptitude Q.1 to Q.10

Q.1 MCQ · 1 MARK

General Aptitude › Reasoning

Courage : Bravery :: Yearning : _____

Select the most appropriate option to complete the analogy.

(A) Longing

(B) Yelling

(C) Yawning

(D) Glaring

Official answer: **A**

SOLUTION

1. Courage and Bravery are synonyms, so the analogy pairs a word with its synonym: Courage : Bravery :: Yearning : _____.
2. Yearning means a strong, deep desire for something.
3. Longing also means a strong desire for something, so it is a synonym of Yearning.
4. Yelling means shouting, Yawning means opening the mouth wide when tired or bored, and Glaring means staring angrily. None of these has the meaning of Yearning.
5. Yelling and Yawning only resemble Yearning in sound or spelling, and Glaring has an unrelated meaning.
6. So the second pair must be synonyms in the same way as the first: Yearning : Longing, which is option (A).

KEY IDEA

First find the relationship in the given pair. Courage and Bravery are synonyms, so choose the synonym of Yearning. Ignore options that only sound or look similar.

WATCH OUT FOR

Picking Yelling or Yawning because they sound like Yearning. An analogy depends on meaning, not on sound or spelling.

Q.2 MCQ · 1 MARK

General Aptitude › Grammar & vocabulary

We _____ tennis in the lawn when it suddenly started to rain.

Select the most appropriate option to complete the above sentence.

(A) have been playing

(B) had been playing

(C) would have been playing

(D) could be playing

Official answer: **B**

SOLUTION

1. The clause "it suddenly started to rain" is in the simple past, so the sentence describes a past situation.
2. The blank names an activity that was already in progress when the rain began and was interrupted by it. This needs a tense that shows an ongoing action before another past event.
3. Option (A) "have been playing" is present perfect continuous. It links the action to the present, which clashes with the past event "started to rain".

4. Option (C) "would have been playing" is a hypothetical or unreal past. The sentence reports a real event, so it does not fit.
5. Option (D) "could be playing" expresses present or future possibility, not a real past action.
6. Option (B) "had been playing" is past perfect continuous. It shows an ongoing action that was under way before the later past event (the rain).
7. So the sentence reads: We had been playing tennis in the lawn when it suddenly started to rain.

KEY IDEA
Use the past perfect continuous (had been + verb-ing) for a longer action that was in progress before a later past event, here the rain.

WATCH OUT FOR
Choosing (A) because "have been playing" sounds natural for an ongoing activity. The present perfect continuous cannot be used when the sentence is anchored in a past event.

Q.3 MCQ · 1 MARK

General Aptitude › Data interpretation

Q.3 A 4 × 4 digital image has pixel intensities (U) as shown in the figure. The number of pixels with $U \leq 4$ is:

0	1	0	2
4	7	3	3
5	5	4	4
6	7	3	2

(A) 3

(B) 8

(C) 11

(D) 9

Official answer: **C**

SOLUTION

1. The condition is $U \leq 4$. Because the inequality is non-strict, a pixel with intensity exactly 4 is counted.
2. Row 1 is 0, 1, 0, 2. All four values are ≤ 4, so this row gives 4.
3. Row 2 is 4, 7, 3, 3. The values 4, 3 and 3 qualify and 7 does not, so this row gives 3.
4. Row 3 is 5, 5, 4, 4. Only the two 4s qualify, so this row gives 2.
5. Row 4 is 6, 7, 3, 2. Only 3 and 2 qualify, so this row gives 2.

6. Total: $4 + 3 + 2 + 2 = 11$.
7. Check by the complement: the pixels with $U > 4$ are 7, 5, 5, 6 and 7, which is 5 pixels. So $16 - 5 = 11$. The answer is option (C).

KEY IDEA

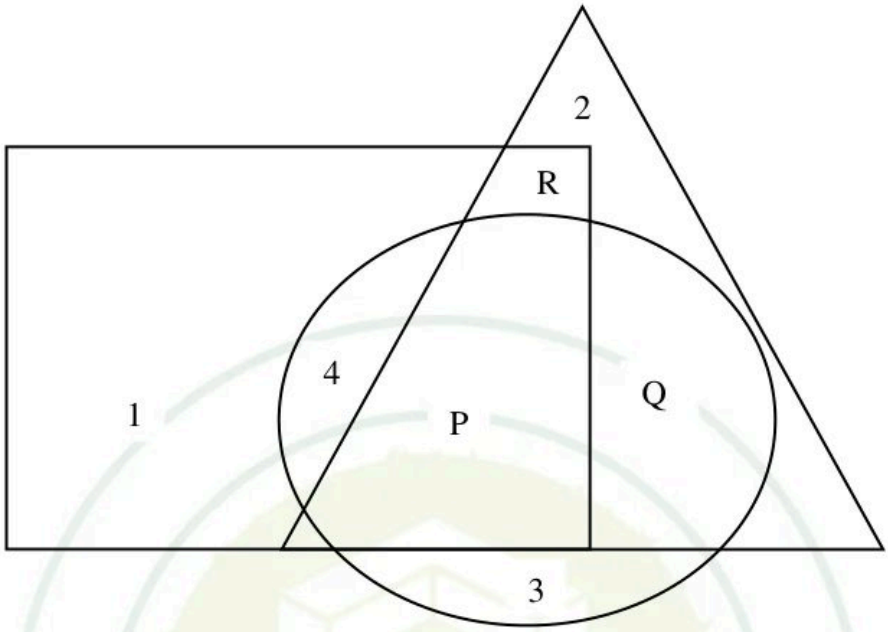
Read the inequality exactly: $U \leq 4$ includes the value 4. Count row by row, then confirm by counting the complement ($U > 4$) and subtracting from 16.

WATCH OUT FOR

Treating $\leq$ as strict $<$ and leaving out the three pixels equal to 4, which gives 8 (option B). Counting the pixels with $U > 4$ (5 of them) and giving that as the answer is another slip.

Q.4 MCQ · 1 MARK

General Aptitude › Reasoning

Q.4	In the given figure, the numbers associated with the rectangle, triangle, and ellipse are 1, 2, and 3, respectively. Which one among the given options is the most appropriate combination of P, Q, and R ?
	
(A)	P = 6; Q = 5; R = 3
(B)	P = 5; Q = 6; R = 3
(C)	P = 3; Q = 6; R = 6
(D)	P = 5; Q = 3; R = 6

Official answer: **A**

SOLUTION

1. Rule from the given labels: each region is labelled with the sum of the numbers of all the shapes that contain it. Rectangle $= 1$, triangle $= 2$, ellipse $= 3$.
2. Check the rule: the region labelled 4 lies inside the rectangle and the ellipse but outside the triangle, so $1 + 3 = 4$. The region labelled 3 is the ellipse alone, and the region labelled 2 is the triangle alone, which agrees with the rule.
3. R lies inside the triangle and the rectangle but above the ellipse, outside it. So $R = 1 + 2 = 3$.

4. Q lies inside the ellipse and the triangle, to the right of the rectangle's right edge. So $Q = 3 + 2 = 5$.
5. P lies inside the rectangle, the triangle and the ellipse. So $P = 1 + 2 + 3 = 6$.
6. Hence $P = 6$, $Q = 5$, $R = 3$, which is option (A).

KEY IDEA
Each region's label is the sum of the numbers of the shapes that contain it. Read which shapes cover a region, then add their numbers.

WATCH OUT FOR
Giving a region only the number of one shape, or missing that Q lies outside the rectangle. This gives Q = 6, as in option (B).

Q.5 MCQ · 1 MARK

General Aptitude › Geometry

A rectangle has a length L and a width W, where L > W. If the width, W, is increased by 10%, which one of the following statements is correct for all values of L and W?

(A) Perimeter increases by 10%.

(B) Length of the diagonals increases by 10%.

(C) Area increases by 10%.

(D) The rectangle becomes a square.

Official answer: **C**

SOLUTION

1. After the change the width is $W' = 1.1W$ and the length L stays the same.
2. Area: $A = LW$ becomes $L(1.1W) = 1.1\,LW$. The area rises by exactly 10% for every L and W, so (C) holds.
3. Perimeter: $2(L + W)$ becomes $2(L + 1.1W)$. The rise is $0.2W$, so the percentage rise is $\dfrac{0.2W}{2(L+W)} \times 100 = \dfrac{10W}{L+W}\%$. Since $L > W$, we have $L + W > 2W$, so this is less than 5%. (A) is false.
4. Diagonal: $d = \sqrt{L^2 + W^2}$ becomes $d' = \sqrt{L^2 + 1.21W^2}$. Only the W^2 part of d^2 grows by 21%, so $d'^2 < 1.21\,d^2$. Hence $d' < 1.1\,d$, and the diagonal grows by less than 10%. (B) is false.
5. Square: the rectangle becomes a square only if $L = 1.1W$. This is not true for all $L > W$, for example $L = 2W$. (D) is false.
6. Only option (C) is correct for all values of L and W.

KEY IDEA
Raising one side of a rectangle by 10% multiplies the area by 1.1 exactly. Perimeter and diagonal are sums or roots of both sides, so they rise by less than 10%.

WATCH OUT FOR
Assuming every measure of the rectangle (perimeter, diagonal) grows by the same 10% as the width. Only a measure that is proportional to W alone, such as the area with L fixed, grows by exactly 10%.

Q.6 MCQ · 2 MARKS

General Aptitude › Grammar & vocabulary

Q.6 Column-I has statements made by Shanthala; and, Column-II has responses given by Kanishk.

Column-I		Column-II	
P.	This house is in a mess.	1.	Alright, I won't bring it up during our conversations.
Q.	I am not happy with the marks given to me.	2.	Well, you can easily look it up.
R.	Politics is a subject I avoid talking about.	3.	No problem, let me clear it up for you.
S.	I don't know what this word means.	4.	Don't worry, I will take it up with your teacher.

Identify the option that has the correct match between Column-I and Column-II.

(A)	P – 2; Q – 3; R – 1; S – 4
(B)	P – 3; Q – 4; R – 1; S – 2
(C)	P – 4; Q – 1; R – 2; S – 3
(D)	P – 1; Q – 2; R – 4; S – 3

Official answer: **B**

SOLUTION

1. Each reply in Column-II uses a phrasal verb with "up". Match each statement to the reply whose phrasal verb fits its situation.
2. P ("This house is in a mess."): "No problem, let me clear it up for you." Here "clear up" means to tidy or put in order. So P – 3.
3. Q ("I am not happy with the marks given to me."): "Don't worry, I will take it up with your teacher." Here "take up with" means to raise a matter with someone. So Q – 4.

4. R ("Politics is a subject I avoid talking about."): "Alright, I won't bring it up during our conversations." Here "bring up" means to introduce a topic. So R – 1.
5. S ("I don't know what this word means."): "Well, you can easily look it up." Here "look up" means to search for a word in a dictionary. So S – 2.
6. The matching is P – 3; Q – 4; R – 1; S – 2, which is option (B).
7. Check the others: (A) pairs P with 2, but "look it up" does not answer a complaint about a messy house. (C) pairs P with 4, but taking a messy house up with a teacher makes no sense. (D) pairs P with 1, but promising not to bring something up does not address a mess. Only (B) fits all four.

KEY IDEA
Match each statement to the reply whose phrasal verb fits the situation: clear up (tidy), take up with (raise a matter with someone), bring up (introduce a topic), look up (search for a word).

WATCH OUT FOR
Matching by the shared word "up" or by the order of the rows instead of by meaning. For example, pairing the unknown word with "take it up with your teacher", or the avoided topic with "clear it up".

Q.7 MCQ · 2 MARKS

General Aptitude › Reasoning

Weight of a person can be expressed as a function of their age. The function usually varies from person to person. Suppose this function is identical for two brothers, and it monotonically increases till the age of 50 years and then it monotonically decreases. Let a_1 and a_2 (in years) denote the ages of the brothers and $a_1 < a_2$.

Which one of the following statements is correct about their age on the day when they attain the same weight?

(A) $a_1 < a_2 < 50$

(B) $a_1 < 50 < a_2$

(C) $50 < a_1 < a_2$

(D) Either $a_1 = 50$ or $a_2 = 50$

Official answer: **B**

SOLUTION

1. Let $W(a)$ be the weight at age a. It increases monotonically up to age 50 and decreases monotonically after age 50, so the maximum weight occurs at $a = 50$.
2. The brothers attain the same weight, so $W(a_1) = W(a_2)$ with $a_1 < a_2$.
3. Case $a_1 < a_2 \leq 50$: both ages lie where W is increasing, so $W(a_1) < W(a_2)$. Equal weights are impossible, so this case cannot occur. This rules out (A).
4. Case $50 \leq a_1 < a_2$: both ages lie where W is decreasing, so $W(a_1) > W(a_2)$. Equal weights are impossible, so this case cannot occur. This rules out (C).
5. Option (D) puts one brother at age 50, where the weight is the maximum. The other age lies strictly on one monotonic side of 50, where the weight is strictly smaller than the maximum. Equal weights are impossible, so (D) fails.
6. Every case with both ages on one side of 50, or with one age equal to 50, is therefore impossible. The ages must lie on opposite sides of the peak, so $a_1 < 50 < a_2$. This is option (B).

KEY IDEA

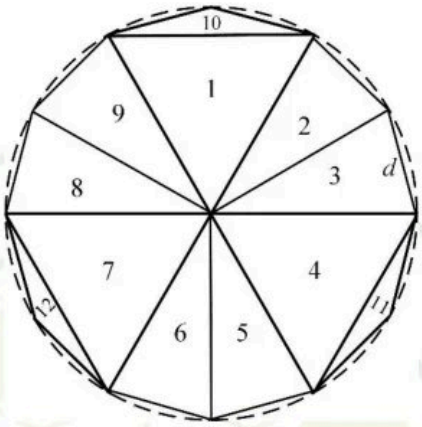
A function that strictly rises up to a peak and then strictly falls takes each value below the peak exactly once on each side of the peak, so two different ages with equal weight must lie on opposite sides of the peak age.

WATCH OUT FOR

Treating "monotonically increasing till 50" as allowing equal weights at two ages on the same side, or choosing (D) because the peak age 50 looks special. At age 50 the weight is the maximum, so no other age can match it.

Q.8 MCQ · 2 MARKS

General Aptitude › Geometry

Q.8	A regular dodecagon (12-sided regular polygon) is inscribed in a circle of radius r cm as shown in the figure. The side of the dodecagon is d cm. All the triangles (numbered 1 to 12) in the figure are used to form squares of side r cm and each numbered triangle is used only once to form a square. The number of squares that can be formed and the number of triangles required to form each square, respectively, are: Note: The figure shown is representative.
	
(A)	3; 4
(B)	4; 3
(C)	3; 3
(D)	3; 2

Official answer: **A**

SOLUTION

1. Joining the centre to the 12 vertices of the regular dodecagon gives 12 congruent isosceles triangles. Each has two sides equal to the radius r.
2. The angle at the centre of each triangle is $\dfrac{360^\circ}{12} = 30^\circ$.
3. Area of one triangle (two sides and the included angle): $\dfrac{1}{2}\, r \cdot r \sin 30^\circ = \dfrac{1}{2}\, r^2 \cdot \dfrac{1}{2} = \dfrac{r^2}{4}$.
4. Area of a square of side r is r^2.

5. Triangles needed per square: $\frac{r^2}{r^2/4} = 4$.
6. Number of squares: $\frac{12}{4} = 3$, since each of the 12 triangles is used once.
7. So 3 squares can be formed, each using 4 triangles: option (A).

KEY IDEA

Each of the 12 triangles has two sides r and an included angle of 30°, so its area is $\frac{1}{2}r^2 \sin 30^\circ = r^2/4$. Divide the square's area r^2 by this to get the triangles per square, and divide 12 by that to get the number of squares.

WATCH OUT FOR

Taking the triangles as equilateral because the sides are called r, or using the side d in the area. The central angle is 30°, not 60°, so the triangle is not equilateral. Also, 12 triangles with 3 per square would give 4 squares, which is the reverse of the correct pairing.

Q.9 MCQ · 2 MARKS

General Aptitude › Numerical

If a real variable x satisfies $3^{x^2} = 27 \times 9^x$, then the value of $\frac{2^{x^2}}{(2^x)^2}$ is:

(A) 2^{-1}
(B) 2^0
(C) 2^3
(D) 2^{15}

Official answer: **C**

SOLUTION

1. Write both sides with base 3: $27 = 3^3$ and $9^x = (3^2)^x = 3^{2x}$.
2. So $3^{x^2} = 3^3 \cdot 3^{2x} = 3^{2x+3}$ (powers of the same base add).
3. Equal bases (with base greater than 1) give equal exponents: $x^2 = 2x + 3$, so $x^2 - 2x - 3 = 0$.
4. Factorise: $(x-3)(x+1) = 0$, so $x = 3$ or $x = -1$.
5. Simplify the required expression: $(2^x)^2 = 2^{2x}$, so $\frac{2^{x^2}}{2^{2x}} = 2^{x^2-2x}$ (dividing powers of the same base subtracts exponents).
6. From the quadratic, $x^2 - 2x = 3$ for both roots, so the value is $2^3 = 8$.
7. Check: $x = 3$ gives $2^{9-6} = 2^3$; $x = -1$ gives $2^{1+2} = 2^3$. The answer is option (C).

KEY IDEA

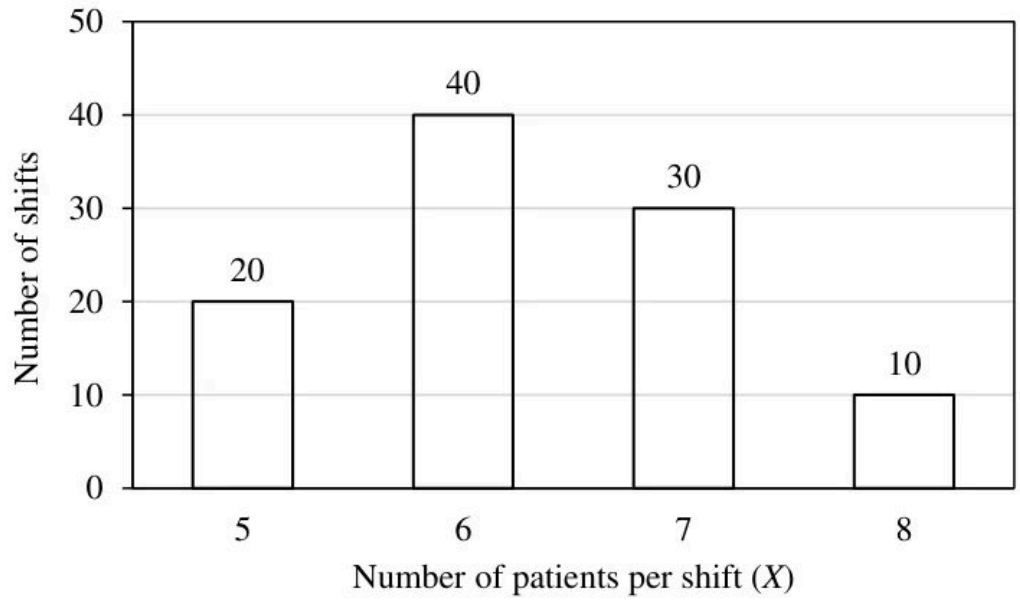
Reduce both sides of the equation to a single base, equate exponents to get $x^2 - 2x = 3$, then notice the asked expression equals 2^{x^2-2x}, so the roots need not be found separately.

WATCH OUT FOR

Reading $(2^x)^2$ as 2^{x^2}, which would make the ratio 1. The power of a power multiplies exponents: $(2^x)^2 = 2^{2x}$.

Q.10 MCQ · 2 MARKS

General Aptitude › Data interpretation

Q.10	The number of patients per shift (X) consulting Dr. Gita in her past 100 shifts is shown in the figure. If the amount she earns is ₹ $1000(X - 0.2)$, what is the average amount (in ₹) she has earned per shift in the past 100 shifts? Note: The figure shown is representative.
	
(A)	6,100
(B)	6,300
(C)	6,000
(D)	6,500

Official answer: **A**

SOLUTION

1. Read the bar chart: $X = 5, 6, 7, 8$ occur in $20, 40, 30, 10$ shifts, and $20 + 40 + 30 + 10 = 100$ shifts, matching the past 100 shifts.
2. The earning $1000(X - 0.2)$ is a linear function of X, so the average earning per shift equals the earning at the average X: $\text{average} = 1000(\bar{X} - 0.2)$.
3. Mean number of patients per shift: $\bar{X} = \dfrac{5(20) + 6(40) + 7(30) + 8(10)}{100}$.

4. Numerator: $100 + 240 + 210 + 80 = 630$, so $\bar{X} = 6.3$.

5. Average amount per shift: $1000(6.3 - 0.2) = 1000 \times 6.1 = 6100$, that is Rs. 6,100, which is option (A).

KEY IDEA

For a linear earning rule, the average earning equals the rule applied to the average value of X. Find the weighted mean of X from the bar heights, then substitute it into 1000(X - 0.2).

WATCH OUT FOR

Using the unweighted mean of 5, 6, 7, 8 (which is 6.5) instead of weighting each value by its number of shifts. That gives Rs. 6,300, which is option (B). The other slip is forgetting to subtract 0.2 and answering 6,300 from the mean 6.3.

Data Science and Artificial Intelligence Q.11 to Q.65

Q.11 MCQ · 1 MARK

Probability and Statistics › Random variables

Suppose X and Y are random variables. The conditional expectation of X given Y is denoted by $E[X \mid Y]$. Then $E[E[X \mid Y]]$ equals

(A) $E[X \mid Y]$

(B) $\dfrac{E[X]}{E[Y]}$

(C) $E[X]$

(D) $E[Y]$

Official answer: **C**

SOLUTION

1. $E[X \mid Y]$ is itself a random variable: it is a function $g(Y)$ of Y, where $g(y) = E[X \mid Y = y]$.
2. Take its expectation over the distribution of Y: $E[E[X \mid Y]] = \sum_y E[X \mid Y = y]\, P(Y = y)$, or $\int g(y) f_Y(y)\, dy$ for continuous Y.
3. Write $E[X \mid Y = y] = \sum_x x\, P(X = x \mid Y = y)$ and substitute: $\sum_y \sum_x x\, P(X = x \mid Y = y)\, P(Y = y)$.
4. Since $P(X = x \mid Y = y)\, P(Y = y) = P(X = x, Y = y)$, this equals $\sum_x x \sum_y P(X = x, Y = y)$.
5. Summing the joint probability over y gives the marginal: $\sum_y P(X = x, Y = y) = P(X = x)$. So the total is $\sum_x x\, P(X = x) = E[X]$.
6. Option (A) is still a random variable, not a number. Options (B) and (D) involve Y's mean, which has no reason to appear. The answer is (C).

KEY IDEA

Law of iterated expectations (tower property): averaging the conditional mean $E[X \mid Y]$ over the distribution of Y gives back the unconditional mean $E[X]$.

WATCH OUT FOR

Treating $E[X \mid Y]$ as if it were already a number and picking (A), or confusing the outer expectation with a ratio or with $E[Y]$. The inner expectation is a function of Y, and the outer one averages it over Y.

Q.12 MCQ · 1 MARK

Linear Algebra › Linear systems

The number of additions and multiplications involved in performing Gaussian elimination on any $n \times n$ upper triangular matrix is of the order

(A) $O(n)$

(B) $O(n^2)$

(C) $O(n^3)$

(D) $O(n^4)$

Official answer: **B**

SOLUTION

1. Gaussian elimination reduces a matrix to upper triangular form by subtracting multiples of the pivot row from the rows below it. In an upper triangular matrix every entry below the diagonal is already 0, so every row multiplier is 0 and the elimination phase needs no row operations.
2. The arithmetic that remains is the solution of $U\mathbf{x} = \mathbf{b}$ by back substitution, computing $x_n, x_{n-1}, \ldots, x_1$ in turn. Row i uses the $n - i$ already-known unknowns: $n - i$ multiplications, $n - i$ subtractions, and one division by u_{ii}.

3. Total multiplications: $\sum_{i=1}^{n}(n-i) = \frac{n(n-1)}{2}$. The number of additions/subtractions is the same, so multiplications plus additions come to $n(n-1)$ (plus n divisions).
4. $n(n-1) = n^2 - n$ grows like n^2, so the work is $\Theta(n^2)$. Every one of the $\frac{n(n+1)}{2}$ nonzero entries is used, so the work cannot be $O(n)$, and it is not of order n^3 or n^4.
5. Hence the order is $O(n^2)$, option (B). The $O(n^3)$ cost belongs to a full (dense) matrix, where each of about n pivots updates an $(n-k) \times (n-k)$ block.

KEY IDEA
An upper triangular matrix needs no elimination. The only work left is back substitution, which costs about $\frac{n(n-1)}{2}$ multiplications and as many additions, so $\Theta(n^2)$.

WATCH OUT FOR
Quoting the dense-matrix cost $O(n^3)$ for Gaussian elimination without noticing that this matrix is already triangular and the elimination cost vanishes.

Q.13 MCQ · 1 MARK

Linear Algebra › Eigenvalues

The sum of the elements in each row of $A \in \mathbb{R}^{n \times n}$ is 1. If $B = A^3 - 2A^2 + A$, which one of the following statements is correct (for $x \in \mathbb{R}^n$)?

(A) The equation $Bx = 0$ has no solution
(B) The equation $Bx = 0$ has exactly two solutions
(C) The equation $Bx = 0$ has infinitely many solutions
(D) The equation $Bx = 0$ has a unique solution

Official answer: **C**

SOLUTION

1. Let $\mathbf{1}$ be the all-ones vector in $\mathbb{R}^n$. Every row of A sums to 1, so the i-th entry of $A\mathbf{1}$ is $\sum_j a_{ij} = 1$. Hence $A\mathbf{1} = \mathbf{1}$, and 1 is an eigenvalue of A with eigenvector $\mathbf{1} \neq \mathbf{0}$.
2. If $A\mathbf{v} = \lambda\mathbf{v}$, then $A^k\mathbf{v} = \lambda^k\mathbf{v}$ for every $k \geq 1$, so $p(A)\mathbf{v} = p(\lambda)\mathbf{v}$ for any polynomial p. Here $p(\lambda) = \lambda^3 - 2\lambda^2 + \lambda = \lambda(\lambda-1)^2$.
3. At $\lambda = 1$: $p(1) = 1 \cdot 0 = 0$. Therefore $B\mathbf{1} = p(1)\mathbf{1} = \mathbf{0}$ with $\mathbf{1} \neq \mathbf{0}$.
4. So B has a nonzero vector in its null space. Hence $\det B = 0$, B is singular, and $\dim \ker B \geq 1$.
5. $x = \mathbf{0}$ always satisfies $Bx = \mathbf{0}$, so (A) is false. The null space is nontrivial, so $x = \mathbf{0}$ is not the only solution and (D) is false.
6. The solutions of $Bx = \mathbf{0}$ form a subspace of $\mathbb{R}^n$. Every scalar multiple $c\mathbf{1}$ with $c \in \mathbb{R}$ is a solution, and these are all different, so there are infinitely many solutions. The number of solutions can never be exactly two, so (B) is false. The answer is (C).

KEY IDEA
A matrix whose rows each sum to 1 has eigenvalue 1 with eigenvector $\mathbf{1}$. The polynomial $\lambda^3 - 2\lambda^2 + \lambda = \lambda(\lambda-1)^2$ vanishes at $\lambda = 1$, so B has a zero eigenvalue. A square matrix with a zero eigenvalue is singular, and a homogeneous system with a singular matrix has infinitely many solutions.

WATCH OUT FOR
Assuming $Bx = 0$ must have only the trivial solution, or that it has no solution. A homogeneous system always has $x = 0$ as a solution. Also, a linear system over the reals never has exactly two solutions: it has 0, 1 or infinitely many.

Q.14 MCQ · 1 MARK

Calculus and Optimization › Taylor series

Let $f(x) = \dfrac{e^x - e^{-x}}{2}$, $x \in \mathbb{R}$. Let $f^{(k)}(a)$ denote the k^{th} derivative of f evaluated at a. What is the value of $f^{(10)}(0)$? (Note: ! denotes factorial)

(A) 0

(B) 1

(C) $\dfrac{1}{10!}$

(D) $\dfrac{2}{10!}$

Official answer: **A**

SOLUTION

1. Recognise $f(x) = \dfrac{e^x - e^{-x}}{2} = \sinh x$.
2. Differentiate once: $f'(x) = \dfrac{e^x + e^{-x}}{2}$. Differentiate again: $f''(x) = \dfrac{e^x - e^{-x}}{2} = f(x)$.
3. Because $f'' = f$, every second derivative returns the function: $f^{(2k)}(x) = f(x)$ for all $k \geq 0$. Taking $k = 5$ gives $f^{(10)}(x) = f(x)$.
4. Evaluate at 0: $f^{(10)}(0) = f(0) = \dfrac{e^0 - e^{-0}}{2} = \dfrac{1-1}{2} = 0$.
5. Check with the Maclaurin series. Using $e^x = \sum_{n\geq 0} \dfrac{x^n}{n!}$ and $e^{-x} = \sum_{n\geq 0} \dfrac{(-1)^n x^n}{n!}$, we get $f(x) = \dfrac{1}{2}\sum_{n\geq 0} \dfrac{(1-(-1)^n)\,x^n}{n!} = x + \dfrac{x^3}{3!} + \dfrac{x^5}{5!} + \cdots$. Only odd powers appear, so the coefficient of x^{10} is 0.
6. Taylor's relation says $f^{(10)}(0) = 10! \times (\text{coefficient of } x^{10}) = 10! \times 0 = 0$, which agrees with step 4. The answer is option (A).

KEY IDEA

$f(x) = \sinh x$ is an odd function, so its Maclaurin series has only odd powers. Every even-order derivative at 0 is therefore 0. Equivalently, $f'' = f$ and $f(0) = 0$.

WATCH OUT FOR

Using $f^{(n)}(0) = n! \times \frac{1}{n!}$ for every n and picking 1, or confusing the coefficient $\frac{1}{10!}$ with the derivative. The even-power coefficients of $\sinh x$ are zero. Another slip is treating $f^{(10)}$ like the derivative of $\cosh x$, whose value at 0 is 1.

Q.15 MCQ · 1 MARK

Artificial Intelligence › Logic

Let p and q be any two propositions. Consider the following propositional statements.

$S_1 : p \to q$, $S_2 : \neg p \wedge q$, $S_3 : \neg p \vee q$, $S_4 : \neg p \vee \neg q$,

where $\wedge$ denotes conjunction (AND operation), $\vee$ denotes disjunction (OR operation), and $\neg$ denotes negation (NOT operation). Which one of the following options is correct?

(Note: $\equiv$ denotes logical equivalence)

(A) $S_1 \equiv S_3$

(B) $S_2 \equiv S_3$

(C) $S_2 \equiv S_4$

(D) $S_1 \equiv S_4$

Official answer: **A**

SOLUTION

1. Implication elimination: $p \to q \equiv \neg p \vee q$. So $S_1 \equiv S_3$, which is option (A).
2. To rule out the other options, evaluate the statements at $p = F,\ q = F$: $S_1 = T$, $S_2 = \neg F \wedge F = F$, $S_3 = \neg F \vee F = T$, $S_4 = \neg F \vee \neg F = T$.
3. (B) $S_2 \equiv S_3$ is false: at $p = F,\ q = F$, $S_2 = F$ but $S_3 = T$.
4. (C) $S_2 \equiv S_4$ is false: at $p = F,\ q = F$, $S_2 = F$ but $S_4 = T$.
5. (D) $S_1 \equiv S_4$ is false: at $p = T,\ q = T$, $S_1 = T$ but $S_4 = \neg T \vee \neg T = F$.

6. Only (A) holds for every assignment, so the answer is (A).

KEY IDEA
An implication is equivalent to a disjunction: $p \to q \equiv \neg p \vee q$. To refute a claimed equivalence, find one assignment where the two statements differ.

WATCH OUT FOR
Confusing $\neg p \vee \neg q$ (which equals $\neg(p \wedge q)$) with $\neg p \vee q$. Another slip is to read $\neg p \wedge q$ as the negation of the implication. The negation of $p \to q$ is $p \wedge \neg q$.

Q.16 MCQ · 1 MARK

Database Management and Warehousing › Normal forms

If a relational decomposition is not dependency-preserving, which one of the following relational operators will be executed more frequently in order to maintain the dependencies?

(A) Selection
(B) Projection
(C) Join
(D) Set union

Official answer: **C**

SOLUTION

1. A decomposition of R into $R_1, \ldots, R_n$ is dependency-preserving if every functional dependency of F can be enforced inside a single decomposed relation, i.e. $(F_1 \cup F_2 \cup \cdots \cup F_n)^+ = F^+$, where F_i is the projection of F onto R_i.
2. If the decomposition is not dependency-preserving, at least one dependency $X \to Y$ has its attributes spread over two or more decomposed relations, so no single table can verify it.
3. To check such a dependency after an insert or update, the attributes of X and Y must first be brought side by side in one relation. The operator that combines tuples of different relations on common attributes is the (natural) join: $R_1 \bowtie R_2 \bowtie \cdots \bowtie R_n$. The dependency is then tested on the joined result.
4. Selection (σ) and projection (π) act on one relation and cannot bring attributes of different relations together. Union ($\cup$) stacks tuples of relations that have the same schema; it does not place attributes of different relations side by side. So none of these helps check a dependency that spans relations.
5. Hence the join is the operator executed more frequently to maintain the dependencies. Answer: (C).

KEY IDEA
Dependency preservation means each functional dependency can be checked inside one decomposed table. If it is not preserved, checking a spanning dependency needs the tables recombined, which is a join.

WATCH OUT FOR
Choosing projection because decomposition itself is done by projection. Projection is used to create the decomposition. Checking a dependency that spans relations needs the pieces joined back together.

Q.17 MCQ · 1 MARK

Database Management and Warehousing › Relational algebra

Consider the following three relations:

Car (model, year, serial, color)
Make (maker, model)
Own (owner, serial)

A tuple in Car represents a specific car of a given model, made in a given year, with a serial number and a color. A tuple in Make specifies that a maker company makes cars of a certain model. A tuple in Own

specifies that an owner owns the car with a given serial number. Keys are underlined; (owner, serial) together form key for Own. ($\bowtie$ denotes natural join)

$\pi_{\text{owner}}(\text{Own} \bowtie (\sigma_{\text{color='red'}}(\text{Car} \bowtie (\sigma_{\text{maker='ABC'}}\text{Make}))))$

Which one of the following options describes what the above expression computes?

(A) All owners of a red car, a car made by ABC, or a red car made by ABC

(B) All owners of more than one car, where at least one car is red and made by ABC

(C) All owners of a red car made by ABC

(D) All red cars made by ABC

Official answer: **C**

SOLUTION

1. Evaluate from the innermost part outward. $\sigma_{\text{maker='ABC'}}(\text{Make})$ keeps only the Make tuples whose maker is ABC. This is a relation $(\text{maker}, \text{model})$ listing the models that ABC makes.
2. $\text{Car} \bowtie (\ldots)$ is a natural join. The only attribute common to $\text{Car}(\text{model}, \text{year}, \text{serial}, \text{color})$ and $(\text{maker}, \text{model})$ is model. The join therefore keeps exactly those cars whose model is made by ABC, and it attaches the maker column to them.
3. $\sigma_{\text{color='red'}}(\ldots)$ is applied to that result, so it keeps only the red cars among the cars made by ABC. The two conditions are applied one after the other, so a car must satisfy both. The condition is an AND, not an OR. The result is the set of red cars made by ABC.
4. $\text{Own} \bowtie (\ldots)$ is a natural join on the only common attribute, serial. Each Own tuple $(\text{owner}, \text{serial})$ is paired with the car that has that serial number. Only owners of red cars made by ABC remain, with one tuple per (owner, car) pair.
5. π_{owner} keeps just the owner column and removes duplicate rows. Each owner therefore appears once, however many such cars they own.
6. So the expression returns the set of owners who own at least one red car made by ABC, which is option (C). Option (D) is wrong because the output has only the owner attribute, so it returns owners, not cars. Option (A) is wrong because it describes an OR of the two conditions, while the expression requires both. Option (B) is wrong because nothing in the expression requires an owner to hold more than one car.

KEY IDEA

Read a nested relational-algebra expression from the inside out. Nested selections compose as an AND, and a natural join pairs tuples on the common attributes (`model` for Car and Make, `serial` for Own and Car). The final projection onto `owner` returns owners, not cars.

WATCH OUT FOR

Choosing (D) because the red-and-ABC cars are what the inner part computes, and forgetting that the outer join with Own and the projection on `owner` turn the result into a set of owners. Another slip is reading the nested selections as an OR, which gives (A).

Q.18 MCQ · 1 MARK

Data Structures and Algorithms › Sorting & hashing

Consider a hash table of size 10 with indices $\{0, 1, \ldots, 9\}$, with the hash function

$$h(x) = 3x \pmod{10},$$

where linear probing is used to handle collisions. The hash table is initially empty and then the following sequence of keys is inserted into the hash table: $1, 4, 5, 6, 14, 15$. The indices where the keys 14 and 15 are stored are, respectively

(A) 2 and 5

(B) 2 and 6

(C) 4 and 5

(D) 4 and 6

Official answer: **D**

SOLUTION

1. Compute each key's home slot with $h(x) = 3x \bmod 10$ and insert in the given order. On a collision, probe forward one slot at a time (linear probing), wrapping around the table.
2. Key 1: $h(1) = 3 \bmod 10 = 3$. Slot 3 is empty, so 1 goes to slot 3.
3. Key 4: $h(4) = 12 \bmod 10 = 2$. Slot 2 is empty, so 4 goes to slot 2.
4. Key 5: $h(5) = 15 \bmod 10 = 5$. Slot 5 is empty, so 5 goes to slot 5.
5. Key 6: $h(6) = 18 \bmod 10 = 8$. Slot 8 is empty, so 6 goes to slot 8.
6. Key 14: $h(14) = 42 \bmod 10 = 2$. Slot 2 is occupied by 4 and slot 3 by 1. Slot 4 is empty, so 14 goes to slot 4.
7. Key 15: $h(15) = 45 \bmod 10 = 5$. Slot 5 is occupied by 5. Slot 6 is empty, so 15 goes to slot 6.
8. Keys 14 and 15 are stored at indices 4 and 6, which is option (D).

KEY IDEA

With linear probing, insert keys in the given order. A key goes to its home slot $h(x)$ if that slot is free. Otherwise it goes to the next free slot, checking $h(x)+1, h(x)+2, \ldots$ modulo the table size.

WATCH OUT FOR

Using $x \bmod 10$ instead of $3x \bmod 10$. Also stopping after one probe: key 14's home slot 2 is full, and slot 3 is full too, so it must go on to slot 4.

Q.19 MCQ · 1 MARK

Probability and Statistics › Continuous distributions

Let X be a continuous random variable whose cumulative distribution function (CDF) $F_X(x)$, for some t, is given as follows:

$$F_X(x) = \begin{cases} 0 & x \le t \\ \dfrac{x-t}{4-t} & t \le x \le 4 \\ 1 & x \ge 4 \end{cases}$$

If the median of X is 3, then what is the value of t?

(A) 2

(B) 1

(C) -1

(D) 0

Official answer: **A**

SOLUTION

1. The median m of a continuous random variable satisfies $F_X(m) = \frac{1}{2}$. Here $m = 3$, so we need $F_X(3) = \frac{1}{2}$.
2. The CDF is 0 for $x \le t$ and 1 for $x \ge 4$. A value of $\frac{1}{2}$ is strictly between 0 and 1, so 3 must lie in the middle piece, $t < 3 < 4$. There $F_X(x) = \dfrac{x-t}{4-t}$.
3. Set $\dfrac{3-t}{4-t} = \dfrac{1}{2}$.
4. Cross-multiply: $2(3-t) = 4-t$, so $6 - 2t = 4 - t$.
5. Rearranging gives $t = 2$, which satisfies $t < 3$, so the assumption in step 2 holds.

6. Check: $F_X(3) = \dfrac{3-2}{4-2} = \dfrac{1}{2}$. Here X is uniform on $[2, 4]$, whose midpoint is 3, as expected. The other options give $F_X(3) = \frac{2}{3}$ for $t = 1$, $\frac{3}{4}$ for $t = 0$ and $\frac{4}{5}$ for $t = -1$, none of which equals $\frac{1}{2}$.

KEY IDEA

For a continuous random variable the median m is the point where the CDF equals $\frac{1}{2}$. Solve $F_X(3) = \frac{1}{2}$ using the piece of the CDF that contains 3.

WATCH OUT FOR

Setting the median to the midpoint of 0 and 4, or using the wrong piece of the CDF. Another slip is cross-multiplying incorrectly and getting $t = 1$ from $3 - t = \frac{1}{2}(4 - t)$ treated as $3 - t = 2 - t$. Always substitute the answer back and check that $F_X(3) = \frac{1}{2}$.

Q.20 MCQ · 1 MARK

Probability and Statistics › Random variables

Let $X = aZ + b$, where Z is a standard normal random variable, and a, b are two unknown constants. It is given that $E[X] = 1$, $E[(X - E[X])Z] = -2$, $E[(X - E[X])^2] = 4$, where $E[X]$ denotes the expectation of random variable X. The values of a, b are:

(A) $a = -2, b = 1$

(B) $a = 2, b = -1$

(C) $a = -2, b = -1$

(D) $a = 1, b = 1$

Official answer: **A**

SOLUTION

1. Since Z is standard normal, $E[Z] = 0$ and $E[Z^2] = \mathrm{Var}(Z) + (E[Z])^2 = 1$.
2. By linearity, $E[X] = aE[Z] + b = b$. Given $E[X] = 1$, we get $b = 1$.
3. Then $X - E[X] = aZ + b - b = aZ$.
4. $E[(X - E[X])Z] = E[aZ^2] = aE[Z^2] = a$. Given this equals -2, we get $a = -2$.
5. Check with the variance condition: $E[(X - E[X])^2] = a^2E[Z^2] = a^2 = (-2)^2 = 4$, which matches the given value. This condition alone only gives $a = \pm 2$, so the sign comes from the covariance condition in the previous step.
6. So $a = -2$, $b = 1$, which is option (A).

KEY IDEA

Write $X - E[X] = aZ$. Then $E[(X - E[X])Z] = a\,E[Z^2] = a$ because $E[Z^2] = 1$ for a standard normal. The mean gives b and the covariance condition gives a with its sign.

WATCH OUT FOR

Using only $E[(X - E[X])^2] = a^2 = 4$ gives $a = \pm 2$, so option (A) and option (B) both look possible. Option (B) has $b = -1$, which contradicts $E[X] = 1$, and $a = 2$ contradicts the covariance condition. Another slip is to write $E[Z^2] = 0$ by confusing it with $E[Z] = 0$.

Q.21 MCQ · 1 MARK

Probability and Statistics › Continuous distributions

It is given that $P(X \geq 2) = 0.25$ for an exponentially distributed random variable X with $E[X] = \dfrac{1}{\lambda}$, where $E[X]$ denotes the expectation of X. What is the value of λ? (ln denotes natural logarithm)

(A) $\ln 2$

(B) $\ln 4$

(C) $\ln 3$

(D) $\ln 0.25$

Official answer: **A**

SOLUTION

1. An exponential random variable with rate λ (so $E[X] = 1/\lambda$) has survival function $P(X \geq x) = e^{-\lambda x}$ for $x \geq 0$. This follows from $P(X \geq x) = \int_x^\infty \lambda e^{-\lambda t}\, dt = e^{-\lambda x}$.
2. Set $x = 2$: $e^{-2\lambda} = 0.25 = \dfrac{1}{4}$.
3. Take natural logs of both sides: $-2\lambda = \ln \dfrac{1}{4} = -\ln 4$.
4. So $\lambda = \dfrac{\ln 4}{2}$.
5. Use $\ln 4 = 2 \ln 2$ to get $\lambda = \dfrac{2 \ln 2}{2} = \ln 2$.
6. Check: $e^{-2 \ln 2} = 2^{-2} = 0.25$, as given. The answer is option (A).

KEY IDEA

For an exponential variable with rate λ, the tail probability is $P(X \geq x) = e^{-\lambda x}$. Set this equal to the given value and take natural logs to solve for λ.

WATCH OUT FOR

Using the mean $1/\lambda$ as if it were the rate, or writing $P(X \geq 2) = 1 - e^{-2\lambda}$ (the CDF) instead of $e^{-2\lambda}$. That gives $e^{-2\lambda} = 0.75$ and $\lambda = \frac{1}{2} \ln \frac{4}{3}$, which is not an option. Choosing $\ln 4$ by forgetting to divide by 2 is also common.

Q.22 MCQ · 1 MARK

Machine Learning › Neural networks

Consider designing a linear classifier $y = \text{sign}(f(x; w, b))$, $f(x; w, b) = w^\top x + b$ on a dataset $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\}$, $x_i \in \mathbb{R}^d$, $y_i \in \{+1, -1\}$, $i = 1, 2, \ldots, N$. Recall that the sign function outputs $+1$ if the argument is positive, and -1 if the argument is non-positive. The parameters w and b are updated as per the following training algorithm: $w_{new} = w_{old} + y_n x_n$, $b_{new} = b_{old} + y_n$ whenever $\text{sign}(f(x_n; w_{old}, b_{old})) \neq y_n$. In other words, whenever the classifier wrongly predicts a sample (x_n, y_n) from the dataset, w_{old} gets updated to w_{new}, and likewise b_{old} gets updated to b_{new}. Consider the case $(x_n, +1)$, $f(x_n; w_{old}, b_{old}) < 0$. Then

(A) $f(x_n; w_{new}, b_{new}) > f(x_n; w_{old}, b_{old})$

(B) $f(x_n; w_{new}, b_{new}) < f(x_n; w_{old}, b_{old})$

(C) $f(x_n; w_{new}, b_{new}) = f(x_n; w_{old}, b_{old})$

(D) $y_n f(x_n; w_{old}, b_{old}) > 1$

Official answer: **A**

SOLUTION

1. The sample has $y_n = +1$ and $f(x_n; w_{old}, b_{old}) < 0$. The sign function then outputs $-1 \neq y_n$, so the sample is misclassified and the update fires: $w_{new} = w_{old} + x_n$ and $b_{new} = b_{old} + 1$.
2. Evaluate the new score at the same point x_n:
$$f(x_n; w_{new}, b_{new}) = (w_{old} + x_n)^\top x_n + b_{old} + 1.$$
3. Expand:
$$f(x_n; w_{new}, b_{new}) = w_{old}^\top x_n + b_{old} + x_n^\top x_n + 1 = f(x_n; w_{old}, b_{old}) + \|x_n\|^2 + 1.$$
4. Since $\|x_n\|^2 \geq 0$, the increment $\|x_n\|^2 + 1 \geq 1 > 0$. The score therefore strictly increases: $f(x_n; w_{new}, b_{new}) > f(x_n; w_{old}, b_{old})$. This is option (A).
5. Options (B) and (C) contradict this strict increase. Option (D) is false, because $y_n f(x_n; w_{old}, b_{old}) = f(x_n; w_{old}, b_{old}) < 0$, which is not greater than 1.

KEY IDEA

A perceptron update on a mistake always moves the score at that sample toward the correct side. For $y_n = +1$ the score changes by $y_n(\|x_n\|^2 + 1) = \|x_n\|^2 + 1 > 0$. In general, the change in $f(x_n)$ is $y_n(\|x_n\|^2 + 1)$.

WATCH OUT FOR

Assuming one update must fix the misclassification, or confusing the increase with $f_{new} > 0$. The score increases by $\|x_n\|^2 + 1$, but the new score may still be negative. Another slip is forgetting the bias update, which contributes the extra $+1$, or choosing (D), which wrongly asserts a margin condition on the old score.

Q.23 MCQ · 1 MARK

Data Structures and Algorithms › Python

Consider the following Python declarations of two lists.

```
A=[1,2,3]
B=[4,5,6]
```

Which one of the following statements results in A= [1, 2, 3, 4, 5, 6]?

(A) A.extend(B)

(B) A.append(B)

(C) A.update(B)

(D) A.insert(B)

Official answer: **A**

SOLUTION

1. Start with `A=[1,2,3]` and `B=[4,5,6]`. The target is the flat list `[1, 2, 3, 4, 5, 6]`.
2. (A) `A.extend(B)` appends every element of the iterable `B` to the end of `A`, in place. `A` becomes `[1, 2, 3, 4, 5, 6]`. This matches.
3. (B) `A.append(B)` adds `B` as ONE element. `A` becomes `[1, 2, 3, [4, 5, 6]]`, a nested list of length 4. This does not match.
4. (C) `A.update(B)` fails because `update` is not a list method (it belongs to dict and set). Python raises `AttributeError`.
5. (D) `A.insert(B)` needs two arguments, an index and an object. Called with one argument it raises `TypeError`.
6. Only (A) gives the flat concatenation.

KEY IDEA

`list.extend(iterable)` adds each element of the iterable to the end of the list, in place. `list.append(x)` adds `x` as a single element.

WATCH OUT FOR

Choosing `A.append(B)`. It looks like it joins the lists, but it nests `B` as one element and gives `[1, 2, 3, [4, 5, 6]]`.

Q.24 MSQ · 1 MARK

Calculus and Optimization › Limits & continuity

Consider two functions $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to (1, \infty)$. Both functions are differentiable at a point c. Which of the following functions is/are ALWAYS differentiable at c? The symbol $\cdot$ denotes product and the symbol $\circ$ denotes composition of functions.

(A) $f \pm g$

(B) $f \cdot g$

(C) $\frac{f}{g}$

(D) $f \circ g + g \circ f$

Official answer: **A, B, C**

SOLUTION

1. Sum, difference and product rules: if f and g are differentiable at c, then so are $f \pm g$ and $f \cdot g$. So (A) and (B) are always differentiable at c.
2. Quotient rule: $\frac{f}{g}$ is differentiable at c when f and g are differentiable at c and $g(c) \neq 0$. Since g takes values in $(1, \infty)$, $g(c) > 1$, so $g(c) \neq 0$ and g never vanishes anywhere. Hence (C) is always differentiable at c.
3. Chain rule: $f \circ g$ is differentiable at c when g is differentiable at c AND f is differentiable at the point $g(c)$. We are told only that f is differentiable at c, not at $g(c)$. Likewise $g \circ f$ needs g differentiable at $f(c)$, which is not given. So (D) can fail, and we need one counterexample.
4. Counterexample for (D): take $c = 0$, $f(x) = |x - 2|$ and $g(x) = 2 + \frac{x}{1 + x^2}$. Here f is differentiable at 0 because $f(x) = 2 - x$ for $x < 2$. Also g is differentiable everywhere (the denominator never vanishes), and since $\left|\frac{x}{1 + x^2}\right| \leq \frac{1}{2}$, we get $g(x) \in [1.5, 2.5] \subset (1, \infty)$. So both functions meet the hypotheses at $c = 0$.
5. Compute the two compositions. First, $g(0) = 2$, so $(f \circ g)(x) = \left|\frac{x}{1 + x^2}\right| = \frac{|x|}{1 + x^2}$. Its right-hand slope at 0 is $+1$ and its left-hand slope is -1, so it is NOT differentiable at 0.
6. Second, near $x = 0$ we have $f(x) = 2 - x$, so $(g \circ f)(x) = g(2 - x)$, which is differentiable at 0 because g is differentiable everywhere.
7. A function that is not differentiable at c, plus one that is differentiable at c, is not differentiable at c (otherwise the difference of the sum and the second function would be differentiable). Hence $f \circ g + g \circ f$ is not differentiable at 0, and (D) is not always differentiable.
8. Therefore the functions that are always differentiable at c are (A), (B) and (C).

KEY IDEA
Sum, product and quotient rules need differentiability only at c (with a non-zero denominator, guaranteed here because $g > 1$). The chain rule for a composition needs the outer function to be differentiable at the inner function's value, which is not given.

WATCH OUT FOR
Assuming a composition of differentiable-at-c functions is differentiable at c. The chain rule requires the outer function to be differentiable at $g(c)$ (or $f(c)$), not at c. Another slip is forgetting that the range $(1, \infty)$ is what keeps the denominator of f/g non-zero.

Q.25 MSQ · 1 MARK

Linear Algebra › Vector spaces

Which of the following statements is/are correct?

(A) $\mathbb{R}^n$ has a unique set of orthonormal basis vectors

(B) $\mathbb{R}^n$ does not have a unique set of orthonormal basis vectors

(C) Linearly independent vectors in $\mathbb{R}^n$ are orthonormal

(D) Orthonormal vectors $\mathbb{R}^n$ are linearly independent

Official answer: **B, D**

SOLUTION

1. Test (A) and (B). In $\mathbb{R}^2$, $\{(1, 0), (0, 1)\}$ is an orthonormal basis. So is $\left\{\frac{1}{\sqrt{2}}(1, 1),\ \frac{1}{\sqrt{2}}(-1, 1)\right\}$: each vector has length $\frac{1}{\sqrt{2}}\sqrt{2} = 1$ and their dot product is $\frac{1}{2}(-1 + 1) = 0$. In general, if $\{e_1, \ldots, e_n\}$ is orthonormal and Q is any orthogonal matrix, then $\{Qe_1, \ldots, Qe_n\}$ is another orthonormal basis,

because $(Qe_i)^\top (Qe_j) = e_i^\top Q^\top Q e_j = \delta_{ij}$. So the orthonormal basis is not unique: (A) is false and (B) is true.

2. Test (C). In $\mathbb{R}^2$ the vectors $(1, 0)$ and $(1, 1)$ are linearly independent, since neither is a multiple of the other. But $(1, 0) \cdot (1, 1) = 1 \neq 0$, so they are not orthogonal, and hence not orthonormal. (C) is false.
3. Test (D). Let $v_1, \ldots, v_k$ be orthonormal, so $v_j^\top v_i = \delta_{ij}$. Suppose $\sum_i c_i v_i = 0$. Multiply on the left by $v_j^\top$ to get $\sum_i c_i \, v_j^\top v_i = c_j = v_j^\top \cdot 0 = 0$. This holds for every j, so all coefficients are zero. Orthonormal vectors are linearly independent: (D) is true.
4. Correct statements: (B) and (D).

KEY IDEA

Orthonormal vectors always satisfy $v_j^\top v_i = \delta_{ij}$, so taking the inner product of any zero combination with v_j forces $c_j = 0$. This gives linear independence. The converse is false, and an orthonormal basis can be rotated by any orthogonal matrix, so it is never unique.

WATCH OUT FOR

Treating linear independence and orthonormality as equivalent (picking C), or assuming the standard basis is the only orthonormal basis of $\mathbb{R}^n$ (picking A).

Q.26 MSQ · 1 MARK

Artificial Intelligence › Uncertainty

Which of the following statements is/are correct in a Bayesian network?

(A) Variable elimination is an approximate inference algorithm

(B) Gibbs sampling is an exact inference algorithm

(C) Variable elimination is used to determine conditional probabilities

(D) Rejection sampling is an approximate inference algorithm

Official answer: **C, D**

SOLUTION

1. Inference in a Bayesian network means computing a posterior such as $P(X \mid \mathbf{e})$ from the joint distribution $\prod_i P(X_i \mid \text{Parents}(X_i))$. Inference algorithms are either exact or approximate.
2. (A) Variable elimination sums out the hidden variables one at a time and reuses intermediate factors. It returns the exact posterior, so it is an exact algorithm, not an approximate one. (A) is false.
3. (B) Gibbs sampling is a Markov chain Monte Carlo method. It repeatedly resamples one non-evidence variable at a time given its Markov blanket. Its estimates converge to the true posterior only in the limit of many samples, so it is approximate. (B) is false.
4. (C) Variable elimination computes $P(Q \mid \mathbf{e}) = \alpha \sum_{\text{hidden}} \prod \text{factors}$. It sums out the hidden variables and then normalises, so it is used to determine conditional probabilities. (C) is true.
5. (D) Rejection sampling draws samples from the prior, discards those that disagree with the evidence, and estimates the query from the samples that remain. It is a sampling method, so it is approximate. (D) is true.
6. Correct statements: (C) and (D).

KEY IDEA

Variable elimination is an exact inference method that answers conditional-probability queries. Sampling methods such as rejection sampling and Gibbs sampling are approximate, because their estimates carry sampling error.

WATCH OUT FOR

Calling variable elimination approximate because it looks like a computational shortcut. It is exact, and it is only the sampling methods (rejection, likelihood weighting, Gibbs) that are approximate. Another slip is to take Gibbs sampling as exact because it is guaranteed to converge in the limit.

Q.27 MSQ · 1 MARK

Data Structures and Algorithms › Sorting & hashing

For which of the following inputs does binary search take time $O(\log n)$ in the worst case?

(A) An array of n integers in any order

(B) A linked list of n integers in any order

(C) An array of n integers in increasing order

(D) A linked list of n integers in increasing order

Official answer: **C**

SOLUTION

1. Binary search needs two things: the data must be sorted, and the middle element must be reachable in $O(1)$ time (random access).
2. (A) An array in any order: the array may be unsorted, so binary search can return a wrong answer. It does not work correctly, so it is not an $O(\log n)$ search.
3. (B) A linked list in any order: the list may be unsorted and has no random access. It fails on both counts.
4. (C) An array in increasing order: it is sorted and indexable. Each step reads the middle element in $O(1)$ and halves the range, so $T(n) = T(n/2) + O(1)$, which gives $T(n) = O(\log n)$ in the worst case. It holds.
5. (D) A linked list in increasing order: the order is fine, but reaching the middle node takes pointer traversals. The first step costs about $n/2$, the next about $n/4$, and so on, so the total is $n/2 + n/4 + \cdots = O(n)$, which is not $O(\log n)$. It fails. Comparisons alone are $O(\log n)$, but the time is dominated by traversal.
6. Only option (C) gives $O(\log n)$ worst-case time.

KEY IDEA
Binary search needs sorted data and O(1) random access to the middle element. Only a sorted array provides both.

WATCH OUT FOR
Choosing (D) because a sorted linked list needs only O(log n) comparisons. Reaching each middle node costs O(n) traversal in total, so the running time is O(n).

Q.28 MSQ · 1 MARK

Linear Algebra › Eigenvalues

Let $A = I_n + xx^\top$, where I_n is the $n \times n$ identity matrix and $x \in \mathbb{R}^n$, $x^\top x = 1$. Which of the following options is/are correct?

(A) Rank of A is n

(B) A is invertible

(C) 0 is an eigenvalue of A

(D) A^{-1} has a negative eigenvalue

Official answer: **A, B**

SOLUTION

1. Since $x^\top x = 1$, we have $(xx^\top)x = x(x^\top x) = x$. So x is an eigenvector of $xx^\top$ with eigenvalue 1.
2. If $x^\top y = 0$, then $(xx^\top)y = x(x^\top y) = 0$. These vectors y form an $(n-1)$-dimensional subspace, so $xx^\top$ has eigenvalue 0 with multiplicity $n-1$. The eigenvalues of $xx^\top$ are therefore 1 (once) and 0 ($n-1$ times).
3. $A = I_n + xx^\top$ has the same eigenvectors, and each eigenvalue is shifted by 1. The eigenvalues of A are 2 (once) and 1 ($n-1$ times).

4. $\det A = 2 \cdot 1^{n-1} = 2 \neq 0$, so A is invertible. Option (B) is correct.
5. A matrix with nonzero determinant has full rank, so $\text{rank}(A) = n$. Option (A) is correct.
6. Every eigenvalue of A is 1 or 2, so 0 is not an eigenvalue. Option (C) is wrong.
7. The eigenvalues of A^{-1} are the reciprocals $\frac{1}{2}$ and 1. All are positive, so A^{-1} has no negative eigenvalue. Option (D) is wrong.

KEY IDEA
A rank-one matrix $xx^\top$ with $x^\top x = 1$ has eigenvalues 1 and 0. Adding I_n shifts these to 2 and 1, so A is positive definite, invertible and of full rank.

WATCH OUT FOR
Forgetting that $xx^\top$ is singular, with $n-1$ zero eigenvalues, and then assuming 0 is also an eigenvalue of A. The identity shifts every eigenvalue by 1. Another slip is to take the eigenvalues of A^{-1} as negative.

Q.29 MSQ · 1 MARK

Data Structures and Algorithms › Sorting & hashing

Suppose that insertion sort is applied to the array $[1, 3, 5, 7, 9, 11, x, 15, 13]$ and it takes exactly two swaps to sort the array. Select all possible values of x.

(A) 10
(B) 12
(C) 14
(D) 16

Official answer: **A, C**

SOLUTION

1. Insertion sort moves each element left by adjacent swaps until it reaches its place. Each adjacent swap removes exactly one inversion, and the sort ends with none. So the number of swaps equals the number of inversions, meaning pairs $i < j$ with $a_i > a_j$. We need exactly 2 inversions.
2. The pair $(15, 13)$ is already an inversion, so it contributes 1. The entries $1, 3, 5, 7, 9, 11$ are increasing and smaller than 15 and 13, so no other pair among the fixed entries is an inversion. We therefore need exactly 1 inversion that involves x.
3. Inversions involving x = (number of earlier elements in $\{1, 3, 5, 7, 9, 11\}$ greater than x) + (number of later elements in $\{15, 13\}$ smaller than x).
4. $x = 10$: only $11 > 10$ is greater among the earlier elements, giving 1. Neither 15 nor 13 is smaller than 10, giving 0. Total with x is 1, so the array has 2 inversions and takes 2 swaps. Valid.
5. $x = 12$: no earlier element exceeds 12, and neither 15 nor 13 is smaller than 12. That gives 0 inversions with x, so the array has 1 inversion and takes 1 swap. Not valid.
6. $x = 14$: no earlier element exceeds 14. Among the later elements, $13 < 14$ counts and $15 < 14$ does not, giving 1. The array has 2 inversions and takes 2 swaps. Valid.
7. $x = 16$: no earlier element exceeds 16, but both $15 < 16$ and $13 < 16$ count, giving 2. The array has 3 inversions and takes 3 swaps. Not valid.
8. Hence the possible values are $x = 10$ and $x = 14$, which are options (A) and (C).

KEY IDEA
With adjacent swaps, insertion sort performs exactly one swap per inversion. Count the inversions in the array and set the total equal to 2.

WATCH OUT FOR
Counting only the inversion that x creates and forgetting that $(15, 13)$ is already an inversion. That leads to wanting 2 inversions with x, which gives 16 instead of 10 and 14.

Q.30 MSQ · 1 MARK

Machine Learning › Clustering

Let C_1 and C_2 be two sets of objects. Let $D(x, y)$ be a measure of dissimilarity between two objects x and y. Consider the following definitions of dissimilarity between C_1 and C_2.

$$\text{DIS-1}(C_1, C_2) = \max_{x \in C_1,\, y \in C_2} D(x, y)$$

$$\text{DIS-2}(C_1, C_2) = \min_{x \in C_1,\, y \in C_2} D(x, y)$$

Which of the following statements is/are correct?

(A) Single Linkage Clustering uses DIS-1

(B) Single Linkage Clustering uses DIS-2

(C) Complete Linkage Clustering uses DIS-2

(D) Complete Linkage Clustering uses DIS-1

Official answer: **B, D**

SOLUTION

1. $\text{DIS-1}(C_1, C_2) = \max_{x \in C_1,\, y \in C_2} D(x, y)$ is the dissimilarity between the two **farthest** points, one from each set.
2. $\text{DIS-2}(C_1, C_2) = \min_{x \in C_1,\, y \in C_2} D(x, y)$ is the dissimilarity between the two **closest** points, one from each set.
3. Single linkage defines the distance between two clusters as the smallest pairwise dissimilarity between them, $\min_{x \in C_1, y \in C_2} D(x, y)$. That is DIS-2, so (B) is correct and (A) is wrong.
4. Complete linkage defines the distance between two clusters as the largest pairwise dissimilarity between them, $\max_{x \in C_1, y \in C_2} D(x, y)$. That is DIS-1, so (D) is correct and (C) is wrong.
5. The correct statements are (B) and (D).

KEY IDEA

Single linkage measures the distance between clusters by their closest pair (a minimum, DIS-2). Complete linkage measures it by their farthest pair (a maximum, DIS-1). Average linkage uses the mean of all pairwise distances.

WATCH OUT FOR

Mixing up the two names, for example reading "single" as "the single farthest point". Single linkage is the nearest-neighbour rule (min) and complete linkage is the farthest-neighbour rule (max).

Q.31 NAT · 1 MARK

Probability and Statistics › Conditional & Bayes

There are three boxes containing white balls and black balls.

Box-1 contains 2 black and 1 white balls.

Box-2 contains 1 black and 2 white balls.

Box-3 contains 3 black and 3 white balls.

In a random experiment, one of these boxes is selected, where the probability of choosing Box-1 is $\frac{1}{2}$, Box-2 is $\frac{1}{6}$, and Box-3 is $\frac{1}{3}$. A ball is drawn at random from the selected box. Given that the ball drawn is white, the probability that it is drawn from Box-2 is _____ (Round off to two decimal places)

Official answer: **0.25**

SOLUTION

1. Let B_i be the event that Box-i is chosen and W the event that the ball drawn is white. The priors are $P(B_1) = \frac{1}{2},\ P(B_2) = \frac{1}{6},\ P(B_3) = \frac{1}{3}$, and they sum to 1.

2. Read the likelihoods from the box contents (white balls out of total balls): $P(W \mid B_1) = \frac{1}{3},\ P(W \mid B_2) = \frac{2}{3},\ P(W \mid B_3) = \frac{3}{6} = \frac{1}{2}$.
3. By the law of total probability:
$$P(W) = \frac{1}{2}\cdot\frac{1}{3} + \frac{1}{6}\cdot\frac{2}{3} + \frac{1}{3}\cdot\frac{1}{2} = \frac{1}{6} + \frac{1}{9} + \frac{1}{6} = \frac{3+2+3}{18} = \frac{4}{9}.$$
4. By Bayes' theorem:
$$P(B_2 \mid W) = \frac{P(W \mid B_2)\,P(B_2)}{P(W)} = \frac{\frac{2}{3}\cdot\frac{1}{6}}{\frac{4}{9}} = \frac{1/9}{4/9} = \frac{1}{4}.$$
5. So the probability is 0.25 to two decimal places.

KEY IDEA
The box is the hidden cause and the white ball is the evidence. Weight each box's chance of giving white by its prior, add these to get $P(W)$, then take Box-2's share of that total (Bayes' theorem).

WATCH OUT FOR
Using the prior $\frac{1}{6}$ directly, or leaving out the prior and comparing only the likelihoods $\frac{1}{3}, \frac{2}{3}, \frac{1}{2}$. That gives $\frac{2/3}{3/2} = \frac{4}{9} \approx 0.44$. The unequal box probabilities must be included in both the numerator and the denominator.

Q.32 NAT · 1 MARK

Calculus and Optimization › Limits & continuity

$\lim\limits_{t\to+\infty} \sqrt{t^2+t} - t =$ _____ (Round off to one decimal place)

Official answer: **0.5**

SOLUTION

1. The limit has the form $\infty - \infty$, so rationalise by multiplying and dividing by the conjugate $\sqrt{t^2+t} + t$.
2. $$\sqrt{t^2+t} - t = \frac{(t^2+t) - t^2}{\sqrt{t^2+t} + t} = \frac{t}{\sqrt{t^2+t} + t}$$
3. For $t > 0$ we have $\sqrt{t^2+t} = t\sqrt{1+\frac{1}{t}}$. Divide the numerator and denominator by t:
$$\frac{t}{\sqrt{t^2+t} + t} = \frac{1}{\sqrt{1+\frac{1}{t}} + 1}$$
4. As $t \to +\infty$, $\frac{1}{t} \to 0$, so $\sqrt{1+\frac{1}{t}} \to 1$.
5. $$\lim_{t\to+\infty} \frac{1}{\sqrt{1+\frac{1}{t}} + 1} = \frac{1}{1+1} = \frac{1}{2} = 0.5$$

KEY IDEA
A difference of two large square-root terms ($\infty - \infty$) is resolved by multiplying by the conjugate, which turns it into a ratio whose limit can be read off after dividing by t.

WATCH OUT FOR
Substituting $\sqrt{t^2+t} \approx t$ and concluding the limit is 0. The neglected term $\frac{1}{2}$ is exactly the answer, since $\sqrt{t^2+t} \approx t + \frac{1}{2}$ for large t.

Q.33 NAT · 1 MARK

Q. 33 On a relation named **Loan** of a bank:

Loan		
loan_number	**branch_name**	**amount**
L11	Banjara Hills	90000
L14	Kondapur	50000
L15	SR Nagar	40000
L22	SR Nagar	25000
L23	Balanagar	80000
L25	Kondapur	70000
L19	SR Nagar	65000

the following SQL query is executed.

```
SELECT L1.loan_number
FROM Loan L1
WHERE L1.amount > (SELECT MAX (L2.amount)
                   FROM Loan L2
                   WHERE L2.branch_name = 'SR Nagar');
```

The number of rows returned by the query is _____(Answer in integer)

Official answer: **3**

SOLUTION

1. The subquery does not refer to `L1`, so it is uncorrelated and is evaluated once. It returns the largest `amount` among the loans whose `branch_name` is 'SR Nagar'.
2. The SR Nagar loans are L15 (40000), L22 (25000) and L19 (65000). So the subquery returns $\text{MAX} = 65000$.
3. The outer query returns every loan, from any branch, whose `amount` is strictly greater than 65000. There is no branch filter in the outer query.
4. Row by row: L11 (90000) passes. L14 (50000) fails. L15 (40000) fails. L22 (25000) fails. L23 (80000) passes. L25 (70000) passes. L19 (65000) fails, because $65000 > 65000$ is false.
5. The qualifying loans are L11, L23 and L25, so the query returns 3 rows.

KEY IDEA
An uncorrelated subquery with an aggregate is evaluated once and gives a single value. Here that value is the SR Nagar maximum, 65000. The outer `>` comparison against it is strict and applies to all branches.

WATCH OUT FOR
Two slips are common. One is to count only SR Nagar rows in the outer query, which has no branch filter. The other is to count L19, whose amount equals the maximum, even though `>` is strict.

Q.34 NAT · 1 MARK

Machine Learning › Regression

Given data $\{(-1, 1), (2, -5), (3, 5)\}$ of the form (x, y), we fit a model $y = wx$ using linear least-squares regression. The optimal value of w is _____. (Round off to three decimal places)

Official answer: **0.285 to 0.287**

SOLUTION

1. The model $y = wx$ has no intercept, so the least-squares objective is $J(w) = \sum_i (y_i - wx_i)^2$, a function of the single parameter w.
2. Setting $\frac{dJ}{dw} = -2\sum_i x_i(y_i - wx_i) = 0$ gives $w\sum_i x_i^2 = \sum_i x_i y_i$. Since $\frac{d^2J}{dw^2} = 2\sum_i x_i^2 > 0$, this stationary point is the minimum. Hence $w = \dfrac{\sum_i x_i y_i}{\sum_i x_i^2}$.
3. Numerator: $\sum x_i y_i = (-1)(1) + (2)(-5) + (3)(5) = -1 - 10 + 15 = 4$.
4. Denominator: $\sum x_i^2 = (-1)^2 + 2^2 + 3^2 = 1 + 4 + 9 = 14$.
5. So $w = \dfrac{4}{14} = \dfrac{2}{7} \approx 0.2857$, which rounds to 0.286.

KEY IDEA
With no intercept, least squares in one parameter gives $w = \dfrac{\sum x_i y_i}{\sum x_i^2}$, which is the projection of the y vector onto the x vector.

WATCH OUT FOR
Fitting a line with an intercept, $y = w_0 + w_1 x$, or using the mean-centred formula $\dfrac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$. The question fixes the model as $y = wx$, so no centring applies.

Q.35 NAT · 1 MARK

Machine Learning › Logistic, k-NN, naive Bayes

The naive Bayes classifier is used to solve a two-class classification problem with class-labels y_1, y_2. Suppose the prior probabilities are $P(y_1) = \frac{1}{3}$ and $P(y_2) = \frac{2}{3}$. Assuming a discrete feature space with

$$P(x \mid y_1) = \frac{3}{4} \quad \text{and} \quad P(x \mid y_2) = \frac{1}{4}$$

for a specific feature vector x. The probability of misclassifying x is _____ (Round off to two decimal places)

Official answer: **0.39 to 0.41**

SOLUTION

1. Naive Bayes assigns x to the class with the larger posterior, $P(y_k \mid x) \propto P(y_k)\, P(x \mid y_k)$.
2. Unnormalised posterior for y_1: $P(y_1)\, P(x \mid y_1) = \frac{1}{3} \cdot \frac{3}{4} = \frac{1}{4}$.
3. Unnormalised posterior for y_2: $P(y_2)\, P(x \mid y_2) = \frac{2}{3} \cdot \frac{1}{4} = \frac{1}{6}$.
4. Evidence by total probability: $P(x) = \frac{1}{4} + \frac{1}{6} = \frac{3}{12} + \frac{2}{12} = \frac{5}{12}$.
5. Posteriors by Bayes' theorem: $P(y_1 \mid x) = \dfrac{1/4}{5/12} = \dfrac{3}{5}$ and $P(y_2 \mid x) = \dfrac{1/6}{5/12} = \dfrac{2}{5}$. They sum to 1, as they should.
6. Since $\frac{3}{5} > \frac{2}{5}$, the classifier predicts y_1.

7. The probability of misclassifying x is the posterior of the class not chosen: $P(y_2 \mid x) = \frac{2}{5} = 0.40$.

KEY IDEA
The error probability of a Bayes classifier at a given x is $1 - \max_k P(y_k \mid x)$. Compute the posteriors by normalising the products of prior and likelihood, then take the smaller one.

WATCH OUT FOR
Using the likelihood $P(x \mid y_2) = \frac{1}{4}$ (or the unnormalised value $\frac{1}{6}$) as the error probability. The likelihoods and the unnormalised posteriors do not sum to 1, so the evidence $P(x)$ must be divided out.

Q.36 MCQ · 2 MARKS

Probability and Statistics › Random variables

Let $Y = Z^2$, $Z = \dfrac{X - \mu}{\sigma}$, where X is a normal random variable with mean μ and variance σ^2. The variance of Y is

(A) 1

(B) 2

(C) 3

(D) 4

Official answer: **B**

SOLUTION

1. Standardising gives $Z = \dfrac{X - \mu}{\sigma} \sim N(0, 1)$, so $Y = Z^2$ is a chi-square variable with 1 degree of freedom. We need $\mathrm{Var}(Y)$.
2. Use $\mathrm{Var}(Y) = E[Y^2] - (E[Y])^2 = E[Z^4] - (E[Z^2])^2$.
3. $E[Z^2] = \mathrm{Var}(Z) + (E[Z])^2 = 1 + 0 = 1$.
4. For the fourth moment, integrate by parts with $\phi'(z) = -z\,\phi(z)$:
$$E[Z^4] = \int_{-\infty}^{\infty} z^3 \cdot z\,\phi(z)\,dz = \Big[-z^3\phi(z)\Big]_{-\infty}^{\infty} + 3\int_{-\infty}^{\infty} z^2\phi(z)\,dz = 0 + 3E[Z^2] = 3.$$
This agrees with the even-moment formula $E[Z^{2k}] = (2k-1)!!$, which gives $3!! = 3$ for $k = 2$.
5. Therefore $\mathrm{Var}(Y) = 3 - 1^2 = 2$. This matches the chi-square variance $2k$ with $k = 1$. The answer is option (B).

KEY IDEA
A standardised normal Z has $E[Z^2] = 1$ and $E[Z^4] = 3$, so $\mathrm{Var}(Z^2) = 3 - 1 = 2$. Equivalently, Z^2 is chi-square with 1 degree of freedom, whose variance is $2k = 2$.

WATCH OUT FOR
Taking $E[Z^4] = 1$ or treating Z^2 like a variable with variance $\mathrm{Var}(Z)^2 = 1$ (option A). The fourth moment of a standard normal is 3, not 1, so $\mathrm{Var}(Z^2) = 3 - 1 = 2$.

Q.37 MCQ · 2 MARKS

Linear Algebra › Eigenvalues

Let $A \in \mathbb{R}^{n \times n}$ be such that $A^3 = A$. Which one of the following statements is ALWAYS correct?

(A) A is invertible

(B) Determinant of A is 0

(C) The sum of the diagonal elements of A is 1

(D) A and A^2 have the same rank

Official answer: **D**

SOLUTION

1. Both $A = I$ and $A = 0$ satisfy $A^3 = A$, so they are valid test matrices for every option.

2. Option (A): for $A = 0$ we have $0^3 = 0$, but 0 is not invertible. So (A) is not always correct.
3. Option (B): for $A = I$ we have $I^3 = I$ and $\det I = 1 \neq 0$. So (B) is not always correct.
4. Option (C): for $A = 0$ the sum of the diagonal elements is $0 \neq 1$, and for $A = I_2$ it is $2 \neq 1$. So (C) is not always correct.
5. Option (D), first inequality: since $A = A^3 = A \cdot A^2$, the column space of A lies inside the column space of A^2 is not needed. We only need $\text{rank}(XY) \leq \min(\text{rank}\, X, \text{rank}\, Y)$. Applied to $A = A \cdot A^2$ it gives $\text{rank}(A) \leq \text{rank}(A^2)$.
6. Option (D), second inequality: applied to $A^2 = A \cdot A$ the same rule gives $\text{rank}(A^2) \leq \text{rank}(A)$.
7. Together, $\text{rank}(A) = \text{rank}(A^2)$ for every real A with $A^3 = A$.
8. Cross-check by eigenvalues: A is annihilated by $x^3 - x = x(x-1)(x+1)$, which has distinct roots. So A is diagonalizable with eigenvalues in $\{0, 1, -1\}$. Its rank is the number of nonzero eigenvalues. A^2 has eigenvalues λ_i^2, and $\lambda_i \neq 0 \iff \lambda_i^2 \neq 0$, so the rank is unchanged.
9. Only (D) holds for every such A.

KEY IDEA

For an ALWAYS-correct statement, test the extreme cases $A = 0$ and $A = I$ to knock out the false options. Then prove the survivor in general. Here $A = A \cdot A^2$ and $A^2 = A \cdot A$ squeeze the ranks together, because a product never has larger rank than either factor.

WATCH OUT FOR

Assuming $A^3 = A$ forces A to be invertible (cancelling A to get $A^2 = I$) or to have a particular determinant or trace. Cancelling A is illegal when A is singular. The matrix 0 and the matrix I both satisfy the condition and give different invertibility, determinant and trace.

Q.38 MCQ · 2 MARKS

Linear Algebra › Special matrices

Let $\{x_1, x_2, \ldots, x_n\}$ be a set of linearly independent vectors in $\mathbb{R}^n$. Let the (i, j)-th element of matrix $A \in \mathbb{R}^{n \times n}$ be given by $A_{ij} = x_i^\top x_j$, $1 \leq i, j \leq n$. Which one of the following statements is correct?

(A) A is invertible

(B) 0 is a singular value of A

(C) Determinant of A is 0

(D) $z^\top A z = 0$ for some non-zero $z \in \mathbb{R}^n$

Official answer: **A**

SOLUTION

1. Let $X = [x_1\ x_2\ \cdots\ x_n] \in \mathbb{R}^{n \times n}$ have the given vectors as its columns. The (i, j) entry of $X^\top X$ is $x_i^\top x_j$, which is A_{ij}. So $A = X^\top X$ (the Gram matrix of the vectors).
2. The n vectors are linearly independent in $\mathbb{R}^n$, so X is square with full column rank. Hence $\text{rank}(X) = n$ and $\det X \neq 0$.
3. Then $\det A = \det(X^\top)\det(X) = (\det X)^2 > 0$. So A is invertible and $\det A \neq 0$. This makes (A) true and (C) false.
4. The singular values of A are the square roots of the eigenvalues of $A^\top A$. Because A is invertible, $A^\top A$ is invertible too, so none of its eigenvalues is 0. Hence 0 is not a singular value of A, and (B) is false.
5. For any non-zero $z \in \mathbb{R}^n$, $z^\top A z = z^\top X^\top X z = \|Xz\|^2$. Since X is invertible, $Xz \neq 0$, so $z^\top A z > 0$. Thus A is positive definite and $z^\top A z = 0$ happens only for $z = 0$. So (D) is false.
6. Only (A) is correct.

KEY IDEA

The matrix with entries $x_i^\top x_j$ is the Gram matrix $A = X^\top X$. When the columns of X are linearly independent, X is invertible, so A is symmetric positive definite and in particular invertible.

WATCH OUT FOR

Treating "Gram matrix" as automatically singular, or mixing up positive semidefinite with positive definite. A Gram matrix is only semidefinite in general. It is singular exactly when the vectors are linearly dependent, and here they are independent.

Q.39 MCQ · 2 MARKS

Probability and Statistics › Continuous distributions

Consider the cumulative distribution function (CDF) of a random variable X:

$$F_X(x) = \begin{cases} 0 & x \le -1 \\ \frac{1}{4}(x+1)^2 & -1 \le x \le 1 \\ 1 & x \ge 1 \end{cases}$$

The value of $P(X^2 \le 0.25)$ is

(A) 0.625

(B) 0.25

(C) 0.5

(D) 0.5625

Official answer: **C**

SOLUTION

1. Rewrite the event: $X^2 \le 0.25 \iff -0.5 \le X \le 0.5$.
2. The CDF is continuous on $[-1, 1]$ (it equals 0 at $x = -1$ and 1 at $x = 1$, and $\frac{1}{4}(x+1)^2$ is continuous), so X has no point masses and $P(X = -0.5) = 0$.
3. Hence $P(-0.5 \le X \le 0.5) = F_X(0.5) - F_X(-0.5)$.
4. $F_X(0.5) = \frac{1}{4}(1.5)^2 = \frac{2.25}{4} = 0.5625$.
5. $F_X(-0.5) = \frac{1}{4}(0.5)^2 = \frac{0.25}{4} = 0.0625$.
6. $P(X^2 \le 0.25) = 0.5625 - 0.0625 = 0.5$, which is option (C).

KEY IDEA

Turn the event on X^2 into an interval for X, then subtract CDF values at the two ends: $P(a \le X \le b) = F(b) - F(a)$ when X has no point mass at a.

WATCH OUT FOR

Using only $F_X(0.5) = 0.5625$ (option D) and forgetting that $X^2 \le 0.25$ also needs $X \ge -0.5$, so the lower tail $F_X(-0.5) = 0.0625$ must be subtracted.

Q.40 MCQ · 2 MARKS

Probability and Statistics › CLT & tests

A random variable X is said to be distributed as Bernoulli(θ), denoted by $X \sim$ Bernoulli(θ), if $P(X = 1) = \theta$, $P(X = 0) = 1 - \theta$ for $0 < \theta < 1$. Let $Y = \sum_{i=1}^{300} X_i$, where $X_i \sim$ Bernoulli(θ), $i = 1, 2, \ldots, 300$ be independent and identically distributed random variables with $\theta = 0.25$. The value of $P(60 \le Y \le 90)$, after approximation through Central Limit Theorem, is given by (Recall that $\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt$)

(A) $\phi(2) - \phi(-2)$

(B) $\phi(1) - \phi(-1)$

(C) $\phi(3) - \phi(-3)$

(D) $\phi(90) - \phi(60)$

Official answer: **A**

SOLUTION

1. Each $X_i \sim \text{Bernoulli}(0.25)$ has $E[X_i] = \theta = 0.25$ and $\text{Var}(X_i) = \theta(1-\theta) = 0.1875$. Since the X_i are independent, $Y = \sum_{i=1}^{300} X_i \sim \text{Binomial}(300, 0.25)$.
2. Mean: $E[Y] = 300 \times 0.25 = 75$.
3. Variance: $\text{Var}(Y) = 300 \times 0.25 \times 0.75 = 56.25$, so the standard deviation is $\sqrt{56.25} = 7.5$.
4. By the Central Limit Theorem, $Z = \dfrac{Y - 75}{7.5}$ is approximately standard normal, so $P(a \le Y \le b) \approx \phi\left(\dfrac{b-75}{7.5}\right) - \phi\left(\dfrac{a-75}{7.5}\right)$, where ϕ is the standard normal CDF as defined in the question.
5. Standardise the limits: $\dfrac{60-75}{7.5} = -2$ and $\dfrac{90-75}{7.5} = 2$.
6. Therefore $P(60 \le Y \le 90) \approx \phi(2) - \phi(-2)$, which is option (A).
7. Option (B) would need limits at 75 ± 7.5 and option (C) at 75 ± 22.5, so neither matches. Option (D) plugs the raw values 60 and 90 into ϕ without standardising.

KEY IDEA

A sum of n i.i.d. Bernoulli(θ) variables has mean $n\theta$ and variance $n\theta(1-\theta)$. The CLT lets you standardise with these and read the probability as a difference of standard normal CDF values.

WATCH OUT FOR

Using the Bernoulli variance $\theta(1-\theta)$ instead of the variance of the sum $n\theta(1-\theta)$. Another slip is forgetting to standardise, which gives option (D), or dividing by the variance 56.25 instead of the standard deviation 7.5.

Q.41 MCQ · 2 MARKS

Probability and Statistics › Continuous distributions

For $x \in \mathbb{R}$, the floor function is denoted by $f(x) = \lfloor x \rfloor$ and defined as follows

$$\lfloor x \rfloor = k, \quad k \le x < k+1,$$

where k is an integer. Let $Y = \lfloor X \rfloor$, where X is an exponentially distributed random variable with mean $\frac{1}{\ln 10}$, where $\ln$ denotes natural logarithm. For any positive integer ℓ, one can write the probability of the event $Y = \ell$ as follows

$$P(Y = \ell) = q^{\ell}(1-q)$$

The value of q is

(A) 0.1

(B) 0.01

(C) 0.5

(D) 0.434

Official answer: **A**

SOLUTION

1. An exponential random variable with mean $\frac{1}{\lambda}$ has rate λ. Here the mean is $\frac{1}{\ln 10}$, so $\lambda = \ln 10$.
2. The event $Y = \ell$ means $\lfloor X \rfloor = \ell$, which holds exactly when $\ell \le X < \ell + 1$.
3. For an exponential variable, $P(X \ge t) = e^{-\lambda t}$ for $t \ge 0$. Hence
$$P(Y = \ell) = P(\ell \le X < \ell + 1) = e^{-\lambda \ell} - e^{-\lambda(\ell+1)}.$$
4. Factor out $e^{-\lambda \ell}$:
$$P(Y = \ell) = e^{-\lambda \ell}\left(1 - e^{-\lambda}\right) = \left(e^{-\lambda}\right)^{\ell}\left(1 - e^{-\lambda}\right).$$
5. Comparing with $q^{\ell}(1-q)$ gives $q = e^{-\lambda}$. So Y is geometric with ratio q.

6. Substitute $\lambda = \ln 10$:

$$q = e^{-\ln 10} = \frac{1}{e^{\ln 10}} = \frac{1}{10} = 0.1.$$

This is option (A).

KEY IDEA

The floor of an exponential variable is geometric. $P(Y = \ell) = P(\ell \le X < \ell + 1) = e^{-\lambda \ell}(1 - e^{-\lambda})$, so the ratio is $q = e^{-\lambda}$, where λ is the rate (the reciprocal of the mean).

WATCH OUT FOR

Taking the stated mean $\frac{1}{\ln 10} \approx 0.434$ as the rate or as q itself (option D). The mean is $\frac{1}{\lambda}$, so $\lambda = \ln 10$ and $q = e^{-\lambda} = 0.1$.

Q.42 MCQ · 2 MARKS

Machine Learning › Neural networks

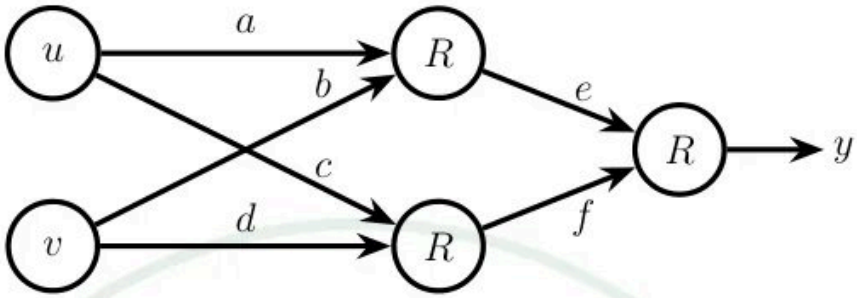
Q. 42 Consider the neural network shown in the figure with

inputs: u, v

weights: a, b, c, d, e, f

output: y

R denotes the ReLU function, $R(x) = \max(0, x)$.



Given $u = 2, v = 3$,

$a = 1, b = 1, c = 1, d = -1, e = 4, f = -1$,

which one of the following is correct?

(A) $\frac{\partial y}{\partial a} = 8, \frac{\partial y}{\partial f} = 0$

(B) $\frac{\partial y}{\partial a} = 1, \frac{\partial y}{\partial f} = 0$

(C) $\frac{\partial y}{\partial a} = 1, \frac{\partial y}{\partial f} = -1$

(D) $\frac{\partial y}{\partial a} = 2, \frac{\partial y}{\partial f} = -1$

Official answer: **A**

SOLUTION

1. Read the figure: a joins u to the top ReLU unit, b joins v to the top unit, c joins u to the bottom unit, and d joins v to the bottom unit. The top and bottom units feed the output ReLU through e and f respectively. So $z_1 = au + bv$, $h_1 = R(z_1)$, $z_2 = cu + dv$, $h_2 = R(z_2)$, $z_3 = eh_1 + fh_2$ and $y = R(z_3)$.
2. Top hidden unit: $z_1 = 1 \cdot 2 + 1 \cdot 3 = 5 > 0$, so $h_1 = R(5) = 5$.
3. Bottom hidden unit: $z_2 = 1 \cdot 2 + (-1) \cdot 3 = -1 < 0$, so $h_2 = R(-1) = 0$.
4. Output unit: $z_3 = 4 \cdot 5 + (-1) \cdot 0 = 20 > 0$, so $y = 20$. The output ReLU is active, with $R'(z_3) = 1$.

5. The derivative of ReLU is $R'(x) = 1$ for $x > 0$ and 0 for $x < 0$. The top unit has $z_1 = 5 > 0$, so $R'(z_1) = 1$.
6. Chain rule for a:
$$\frac{\partial y}{\partial a} = R'(z_3) \cdot e \cdot R'(z_1) \cdot u = 1 \cdot 4 \cdot 1 \cdot 2 = 8.$$
7. Chain rule for f:
$$\frac{\partial y}{\partial f} = R'(z_3) \cdot h_2 = 1 \cdot 0 = 0.$$
The bottom unit is inactive, so $h_2 = 0$ and f has no effect on y at this input.
8. So $\partial y/\partial a = 8$ and $\partial y/\partial f = 0$, which is option (A).

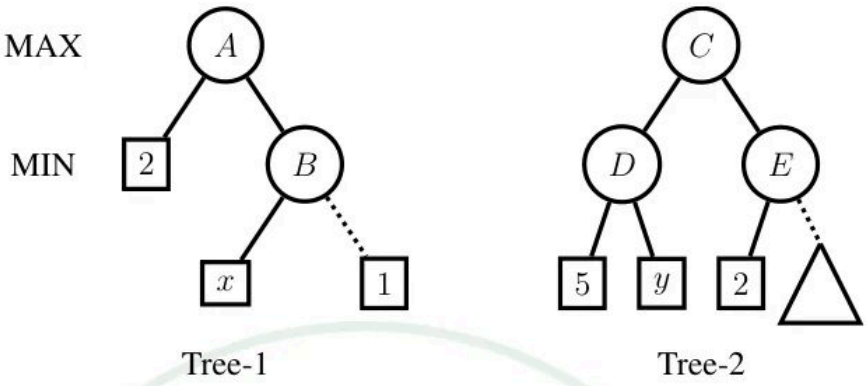
KEY IDEA
Run the forward pass to see which ReLUs are active. Then apply the chain rule through the active path only. A ReLU with a negative pre-activation has slope 0 and outputs 0, so it blocks the gradient and contributes nothing to the next layer.

WATCH OUT FOR
Forgetting the ReLU slopes and taking $\partial y/\partial f = h_2$ or -1 (the weight's value) instead of the bottom unit's output, which is 0. Another slip is to give $\partial y/\partial a = u = 2$ or 1 by leaving out the factor $e = 4$ from the downstream weight.

Q.43 MCQ · 2 MARKS

Artificial Intelligence › Search

Q. 43 Consider game trees Tree-1 and Tree-2 as shown. The first level is a MAX agent and the second level is a MIN agent. The value in the square node is the output of the utility function.



Tree-1 Tree-2

For what ranges of x and y, the right child of node B and the right child of node E will be pruned by alpha-beta pruning algorithm?

(A) $x \in [1, \infty)$ and $y \in (-\infty, 2]$

(B) $x \in (-\infty, 2]$ and $y \in (-\infty, 5]$

(C) $x \in (-\infty, 2]$ and $y \in [2, \infty)$

(D) $x \in [1, \infty)$ and $y \in (-\infty, 5]$

Official answer: **C**

SOLUTION

1. Pruning rule: at a MIN node, once its running value is $\leq \alpha$ (the best value MAX can already guarantee on the path to the root), the remaining children are pruned, because MAX would never choose this

node.

2. Tree-1: A is MAX. Its left child is the leaf 2, so after the left branch $\alpha = 2$ at A.
3. B is MIN and its left leaf is x, so B's running value is at most x. The right child of B (the leaf 1) is pruned exactly when $x \leq \alpha = 2$, that is $x \in (-\infty, 2]$.
4. Tree-2: C is MAX. Its left child D is MIN with leaves 5 and y, so $D = \min(5, y)$ and C gets $\alpha = \min(5, y)$.
5. E is MIN. Its first leaf is 2, so E's running value is at most 2. The right child of E is pruned exactly when $2 \leq \alpha = \min(5, y)$.
6. $2 \leq \min(5, y)$ holds only if both $2 \leq 5$ (always true) and $2 \leq y$. So $y \in [2, \infty)$.
7. Both pruning conditions must hold together: $x \in (-\infty, 2]$ and $y \in [2, \infty)$, which is option (C).

KEY IDEA

A MIN node stops searching its remaining children as soon as its running value drops to α or below, where α is the best value MAX has already secured from earlier siblings. Pruning therefore needs the first child to be $\leq \alpha$.

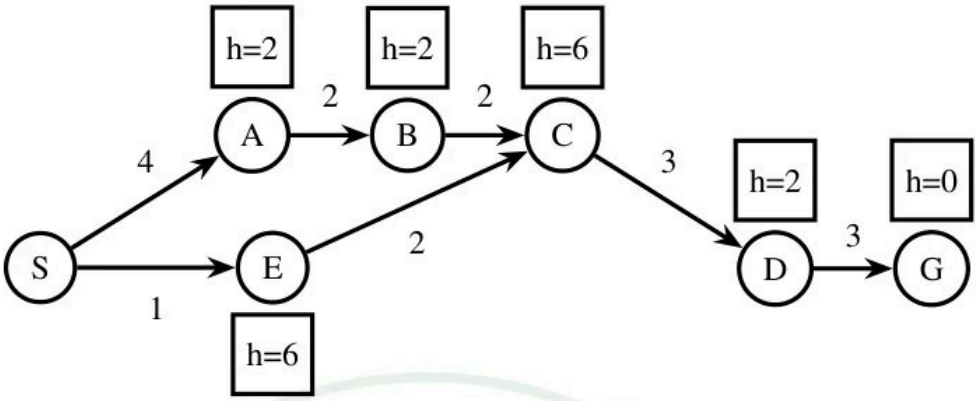
WATCH OUT FOR

Comparing x with the pruned leaf 1 instead of with $\alpha = 2$, which gives $x \geq 1$. Another slip is forgetting that α in Tree-2 is $\min(5, y)$, not y, and writing $y \leq 5$ or $y \leq 2$.

Q.44 MCQ · 2 MARKS

Artificial Intelligence › Search

Q. 44 The state graph shows the action cost along the edges and the heuristic function h associated with each state.



Suppose A^* algorithm is applied on this state graph using priority queue to store the frontier. In what sequence are the nodes expanded?

(A) S,A,E,C,B,D,G

(B) S,E,A,C,B,D,G

(C) S,A,E,B,C,D,G

(D) S,A,B,E,C,D,G

Official answer: **C**

SOLUTION

1. A* always expands the frontier node with the smallest $f(n) = g(n) + h(n)$, where $g(n)$ is the cheapest path cost found so far from S and $h(n)$ is the printed heuristic. Edge costs: $S \to A = 4$, $S \to E = 1$, $A \to B = 2$, $B \to C = 2$, $E \to C = 2$, $C \to D = 3$, $D \to G = 3$.

2. Expand S. Its successors enter the frontier: A with $g = 4$, $f = 4 + 2 = 6$, and E with $g = 1$, $f = 1 + 6 = 7$.
3. The smallest f is A (6), so expand A. Push B with $g = 4 + 2 = 6$, $f = 6 + 2 = 8$. Frontier: $E(7)$, $B(8)$.
4. The smallest f is E (7), so expand E. Push C with $g = 1 + 2 = 3$, $f = 3 + 6 = 9$. Frontier: $B(8)$, $C(9)$.
5. The smallest f is B (8), so expand B. It reaches C with $g = 6 + 2 = 8$, $f = 8 + 6 = 14$. This is worse than the existing entry for C ($g = 3$, $f = 9$), so the entry is kept. Frontier: $C(9)$.
6. Expand C. Push D with $g = 3 + 3 = 6$, $f = 6 + 2 = 8$.
7. Expand D (8). Push G with $g = 6 + 3 = 9$, $f = 9 + 0 = 9$.
8. Expand G, the goal. The expansion order is S, A, E, B, C, D, G, which is option (C). No ties occurred, so no tie-break rule was needed.

KEY IDEA

A* orders the frontier by $f(n) = g(n) + h(n)$, not by g alone and not by h alone. Keep, for each node, the cheapest g found so far, and discard a later path that is costlier.

WATCH OUT FOR

Expanding E before A because E is cheaper to reach ($g = 1$ against 4). That ignores h: $f(E) = 1 + 6 = 7$ is larger than $f(A) = 4 + 2 = 6$. Another slip is expanding C before B. C is reached early through E, but its $f = 9$ is larger than $f(B) = 8$.

Q.45 MCQ · 2 MARKS

Probability and Statistics › Counting & axioms

A random experiment consists of throwing 100 fair dice, each die having six faces numbered 1 to 6. An event A represents the set of all outcomes where at least one of the dice shows a 1. Then, $P(A) =$

(A) 0

(B) 1

(C) $1 - \left(\frac{5}{6}\right)^{100}$

(D) $\left(\frac{5}{6}\right)^{100}$

Official answer: **C**

SOLUTION

1. The event A is "at least one of the 100 dice shows a 1". Its complement A^c is "none of the 100 dice shows a 1", so $P(A) = 1 - P(A^c)$.
2. For one fair die, the faces 2, 3, 4, 5, 6 are not a 1, so $P(\text{not } 1) = \frac{5}{6}$.
3. The 100 dice are thrown independently, so the probabilities multiply:
$$P(A^c) = \left(\frac{5}{6}\right)^{100}.$$
Equivalently, there are 5^{100} outcomes with no 1 out of 6^{100} equally likely outcomes.
4. Therefore
$$P(A) = 1 - \left(\frac{5}{6}\right)^{100}.$$
5. This value is strictly between 0 and 1, so options (A) and (B) are out. Option (D) is $P(A^c)$, not $P(A)$. The answer is (C).

KEY IDEA

For an "at least one" event over independent trials, use the complement: P(at least one) = 1 − P(none), and P(none) is the product of the per-trial "none" probabilities.

WATCH OUT FOR

Choosing (D), which is the probability that no die shows a 1 (the complement), or trying to add probabilities such as 100 × 1/6, which overcounts outcomes where several dice show a 1.

Q.46 MCQ · 2 MARKS

Database Management and Warehousing › Data warehousing

Consider a fact table in an OLAP application: `Facts(D1, D2, val)`, where `D1` and `D2` are its dimension attributes and `val` is a dependent attribute. Suppose attribute `D1` takes 3 values and `D2` takes 2 values, and all combinations of these values are present in the table `Facts`. How many tuples are there in the result of the following query?

```
SELECT D1, D2, sum(val)
FROM Facts
GROUP BY CUBE (D1, D2);
```

(A) 1

(B) 6

(C) 9

(D) 12

Official answer: **D**

SOLUTION

1. `GROUP BY CUBE (D1, D2)` produces the aggregate for every subset of the grouping attributes. With two attributes that is $2^2 = 4$ groupings: $(D1, D2)$, $(D1)$, $(D2)$ and $()$. The result is the union of these four group-bys.
2. Grouping by $(D1, D2)$: every one of the $3 \times 2 = 6$ combinations is present in `Facts`, so this gives 6 tuples.
3. Grouping by $(D1)$ alone, with `D2` shown as NULL: one tuple per value of `D1`, so 3 tuples.
4. Grouping by $(D2)$ alone, with `D1` shown as NULL: one tuple per value of `D2`, so 2 tuples.
5. The empty grouping $()$ is the grand total: 1 tuple with both `D1` and `D2` NULL.
6. Total tuples:
$$6 + 3 + 2 + 1 = 12 = (3 + 1)(2 + 1).$$
7. When every combination is present, a CUBE over attributes with $n_1, n_2, \ldots$ distinct values has $\prod_i (n_i + 1)$ tuples. Here that is $4 \times 3 = 12$, which is option (D).

KEY IDEA

`CUBE` over k attributes is the union of all 2^k group-bys. When every combination of values is present, its size is $\prod_i (n_i + 1)$, because each attribute is either kept (n_i values) or rolled up into NULL (1 extra).

WATCH OUT FOR

Counting only the finest grouping ($3 \times 2 = 6$) as in a plain `GROUP BY`. Another slip is to forget the grand-total row, or to write $2^2 \times$ something instead of adding the four group-by sizes (6 + 3 + 2 + 1).

Q.47 MCQ · 2 MARKS

Data Structures and Algorithms › Python

Consider the following Python code snippet.

```
A={"this","that"}
B={"that","other"}
C={"other","this"}
while "other" in C:
  if "this" in A:
    A,B,C=A-B,B-C,C-A
  if "that" in B:
    A,B,C=C|A,A|B,B|C
```

When the above program is executed, at the end, which of the following sets contains "this"?

(A) Only A

(B) Only B

(C) Only C

(D) A, C

Official answer: **B**

SOLUTION

1. Start: $A = \{\text{this}, \text{that}\}$, $B = \{\text{that}, \text{other}\}$, $C = \{\text{other}, \text{this}\}$. In `A,B,C = X,Y,Z` the whole right side is evaluated first using the old values, then all three names are rebound together.
2. Pass 1, first `if`: "this" is in `A`, so the right side is `A-B` $= \{\text{this}\}$, `B-C` $= \{\text{that}\}$, `C-A` $= \{\text{other}\}$. Now $A = \{\text{this}\}$, $B = \{\text{that}\}$, $C = \{\text{other}\}$.
3. Pass 1, second `if`: "that" is in `B`, so the right side is `C|A` $= \{\text{other}, \text{this}\}$, `A|B` $= \{\text{this}, \text{that}\}$, `B|C` $= \{\text{that}, \text{other}\}$. Now $A = \{\text{other}, \text{this}\}$, $B = \{\text{this}, \text{that}\}$, $C = \{\text{that}, \text{other}\}$.
4. Loop test: "other" is in `C`, so pass 2 runs. First `if`: "this" is in `A`, so `A-B` $= \{\text{other}\}$, `B-C` $= \{\text{this}\}$, `C-A` $= \{\text{that}\}$. Now $A = \{\text{other}\}$, $B = \{\text{this}\}$, $C = \{\text{that}\}$.
5. Pass 2, second `if`: "that" is not in `B` (which is $\{\text{this}\}$), so nothing changes.
6. Loop test: "other" is not in `C` (which is $\{\text{that}\}$), so the loop ends.
7. Final sets: $A = \{\text{other}\}$, $B = \{\text{this}\}$, $C = \{\text{that}\}$. Only `B` contains "this", which is option (B).

KEY IDEA

In a tuple assignment, every expression on the right is computed from the old values before any name is rebound. Trace the three sets pass by pass and re-test each `if` and the `while` condition against the current sets.

WATCH OUT FOR

Forgetting that the second `if` is tested against the already-updated `B`. After pass 2's first `if`, `B` is {this}, so "that" is no longer in it and the second assignment is skipped. Another slip is updating the names one at a time instead of all together.

Q.48 MSQ · 2 MARKS

Machine Learning › Neural networks

Which of the following statements is/are correct about the rectified linear unit (ReLU) activation function defined as $\text{ReLU}(x) = \max(x, 0)$, where $x \in \mathbb{R}$?

(A) ReLU is continuous everywhere

(B) ReLU is differentiable everywhere

(C) ReLU is not differentiable at $x = 0$

(D) $\text{ReLU}(x) = \text{ReLU}(ax)$, for all $a \in \mathbb{R}$

Official answer: **A, C**

SOLUTION

1. Write ReLU piecewise: $\text{ReLU}(x) = 0$ for $x \leq 0$ and $\text{ReLU}(x) = x$ for $x > 0$.
2. Continuity: both pieces are linear, so the only point to check is $x = 0$. The left limit is 0, the right limit is 0, and $\text{ReLU}(0) = 0$. So ReLU is continuous everywhere, and (A) is correct.
3. Differentiability at $x = 0$: the left derivative is $\lim_{h \to 0^-} \frac{0-0}{h} = 0$. The right derivative is $\lim_{h \to 0^+} \frac{h-0}{h} = 1$. They differ, so the derivative does not exist at 0 (a corner).
4. Hence (B) is false, since ReLU is not differentiable at 0, and (C) is true.
5. Test (D) with a negative scale: take $a = -1$ and $x = 1$. Then $\text{ReLU}(1) = 1$ but $\text{ReLU}(-1) = 0$, so the equality fails and (D) is false. It holds only for $a \geq 0$ (positive homogeneity), not for every real a.
6. Correct options: (A) and (C).

KEY IDEA
ReLU is continuous everywhere, but it has a corner at 0 where the left slope is 0 and the right slope is 1, so it is not differentiable there. It is positively homogeneous, so $\mathrm{ReLU}(ax) = a\,\mathrm{ReLU}(x)$ only for $a \geq 0$.

WATCH OUT FOR
Treating ReLU as differentiable everywhere because it is piecewise linear, or assuming $\mathrm{ReLU}(ax) = \mathrm{ReLU}(x)$ for every real a. Negative a flips the sign of the input and changes the output.

Q.49 MSQ · 2 MARKS

Calculus and Optimization › Maxima & minima

Consider the function $f(x) = \frac{x^3}{3} + \frac{7}{2}x^2 + 10x + \frac{133}{2}$, $x \in [-8, 0]$. Which of the following statements is/are correct?

(A) The maximum value of f is attained at $x = -5$

(B) The minimum value of f is attained at $x = -2$

(C) The maximum value of f is $\frac{133}{2}$

(D) The minimum value of the derivative of f is attained at $x = -\frac{7}{2}$

Official answer: **C, D**

SOLUTION

1. Differentiate: $f'(x) = x^2 + 7x + 10 = (x+2)(x+5)$, so the critical points are $x = -5$ and $x = -2$, both inside $[-8, 0]$.
2. Sign of f': positive on $[-8, -5)$, negative on $(-5, -2)$, positive on $(-2, 0]$. So $x = -5$ is a local maximum and $x = -2$ is a local minimum.
3. On a closed interval the absolute extremes occur at a critical point or at an endpoint, so compare $f(-8)$, $f(-5)$, $f(-2)$ and $f(0)$.
4. $f(-8) = -\frac{512}{3} + 224 - 80 + \frac{133}{2} = \frac{239}{6} \approx 39.83$.
5. $f(-5) = -\frac{125}{3} + \frac{175}{2} - 50 + \frac{133}{2} = \frac{187}{3} \approx 62.33$.
6. $f(-2) = -\frac{8}{3} + 14 - 20 + \frac{133}{2} = \frac{347}{6} \approx 57.83$, and $f(0) = \frac{133}{2} = 66.5$.
7. The largest value is $f(0) = \frac{133}{2}$, attained at the endpoint $x = 0$, not at $x = -5$ (where $f \approx 62.33$). So (A) is false and (C) is true.
8. The smallest value is $f(-8) = \frac{239}{6}$, attained at the endpoint $x = -8$, not at $x = -2$ (where $f \approx 57.83$). So (B) is false.
9. For (D), $f'(x) = x^2 + 7x + 10$ is an upward-opening parabola, so it is minimised where its derivative $f''(x) = 2x + 7$ is zero, i.e. $x = -\frac{7}{2}$. This lies in $[-8, 0]$, so (D) is true.
10. Correct options: (C) and (D).

KEY IDEA
On a closed interval the absolute maximum and minimum are found by comparing the function at the critical points and at both endpoints. A local maximum or minimum is not automatically the absolute one. The derivative's own minimum comes from setting the second derivative to zero.

WATCH OUT FOR
Stopping at the critical points $x = -5$ and $x = -2$ and calling them the maximum and minimum, which gives (A) and (B). This ignores the endpoints $x = 0$ and $x = -8$, where the function is actually largest and smallest.

Q.50 MSQ · 2 MARKS

Linear Algebra › Special matrices

Let x_1, x_2, x_3, x_4, x_5 be a system of orthonormal vectors in $\mathbb{R}^{10}$. Consider the matrix $A = x_1 x_1^\top + \ldots + x_5 x_5^\top$. Which of the following statements is/are correct?

(A) Singular values of A are also its eigenvalues

(B) Singular values of A are either 0 or 1

(C) Determinant of A is 1

(D) A is invertible

Official answer: **A, B**

SOLUTION

1. Let $X = [x_1\ x_2\ x_3\ x_4\ x_5]$, a 10×5 matrix. Then $A = \sum_{i=1}^{5} x_i x_i^\top = XX^\top$. Because the x_i are orthonormal, $X^\top X = I_5$.
2. Hence $A^\top = A$ and $A^2 = X(X^\top X)X^\top = XX^\top = A$. So A is a symmetric idempotent matrix, that is, the orthogonal projection onto $\text{span}\{x_1, \ldots, x_5\}$.
3. For $i \le 5$, $Ax_i = \sum_j x_j(x_j^\top x_i) = x_i$, so 1 is an eigenvalue with multiplicity at least 5. Any vector v orthogonal to all x_i satisfies $Av = 0$, and this orthogonal complement has dimension 5. So the eigenvalues are exactly five 1's and five 0's.
4. A is symmetric positive semidefinite, so its singular values are its eigenvalues: five 1's and five 0's. Statement (A) is true, and statement (B) is true because every singular value is 0 or 1.
5. $\det A = 1^5 \cdot 0^5 = 0$, so (C) is false.
6. $\text{rank}\, A = 5 < 10$, so A is singular and not invertible. Statement (D) is false.
7. Correct options: (A) and (B).

KEY IDEA

A sum of $x_i x_i^\top$ over orthonormal vectors is the orthogonal projection onto their span. It is symmetric and idempotent, so its eigenvalues are 0 or 1. Being positive semidefinite, its singular values equal its eigenvalues.

WATCH OUT FOR

Assuming the matrix is invertible or has determinant 1 because the vectors are orthonormal. That holds only when A is built from a full set of 10 orthonormal vectors. With only 5 vectors in $\mathbb{R}^{10}$ the rank is 5, so A is singular and $\det A = 0$.

Q.51 MSQ · 2 MARKS

Calculus and Optimization › Maxima & minima

Let $f : \mathbb{R} \to \mathbb{R}$ be a twice-differentiable function and suppose its second derivative satisfies $f''(x) > 0$ for all $x \in \mathbb{R}$. Which of the following statements is/are ALWAYS correct?

(A) f has a local minima

(B) There does not exist x and y, $x \neq y$, such that $f'(x) = f'(y) = 0$

(C) f has at most one global minimum

(D) f has at most one local minimum

Official answer: **B, C, D**

SOLUTION

1. Since $f''(x) > 0$ for all x, the derivative f' is strictly increasing, so f is strictly convex.
2. (A) Not always. Take $f(x) = e^x$. Then $f''(x) = e^x > 0$ everywhere, but $f'(x) = e^x > 0$ never vanishes, so f has no stationary point and hence no local minimum. False.
3. (B) A strictly increasing function is one-to-one, so $f'(x) = f'(y)$ forces $x = y$. Hence two distinct points $x \neq y$ with $f'(x) = f'(y) = 0$ cannot exist. True.
4. (D) f is differentiable, so at any local minimum c we need $f'(c) = 0$. By (B) there is at most one such point, so there is at most one local minimum. True.
5. (C) Suppose f attained its global minimum value m at two points $a \neq b$, so $f(a) = f(b) = m$. Strict convexity gives $f\left(\frac{a+b}{2}\right) < \frac{f(a)+f(b)}{2} = m$. This contradicts m being the global minimum. So there is at most one global minimum point. True.

6. The statements that always hold are (B), (C) and (D).

KEY IDEA
$f'' > 0$ everywhere makes f' strictly increasing, so f' has at most one zero. That gives at most one stationary point, hence at most one local (and global) minimum. It does not guarantee that a minimum exists.

WATCH OUT FOR
Choosing (A) because a convex function "looks like a bowl". Strict convexity gives uniqueness of a minimum if one exists, not existence. e^x is strictly convex yet has no minimum.

Q.52 MSQ · 2 MARKS

Linear Algebra › Special matrices

An $n \times n$ matrix A with real entries satisfies the property: $\|Ax\|^2 = \|x\|^2$, for all $x \in \mathbb{R}^n$, where $\|\cdot\|$ denotes the Euclidean norm. Which of the following statements is/are ALWAYS correct?

(A) A must be orthogonal

(B) $A = I$, where I denotes the identity matrix, is the only solution

(C) The eigenvalues of A are either $+1$ or -1

(D) A has full rank

Official answer: **A, D**

SOLUTION

1. Expand the condition: $\|Ax\|^2 = (Ax)^T(Ax) = x^T A^T A x$, so the given property says $x^T(A^TA - I)\,x = 0$ for every $x \in \mathbb{R}^n$.
2. Let $M = A^TA - I$. It is symmetric, since $(A^TA)^T = A^TA$. Take $x = e_i$ (a standard basis vector): $e_i^T M e_i = M_{ii} = 0$. Take $x = e_i + e_j$: $0 = (e_i + e_j)^T M (e_i + e_j) = M_{ii} + M_{jj} + 2M_{ij} = 2M_{ij}$, so $M_{ij} = 0$. Hence $M = 0$, that is $A^TA = I$.
3. $A^TA = I$ is the definition of an orthogonal matrix, so (A) is always correct.
4. (B) is false. Every orthogonal matrix satisfies the property, not only I. For example $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ has $A^TA = I$ but $A \neq I$ (a rotation by 90°); so does the reflection $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.
5. (C) is false. The rotation above has characteristic equation $\lambda^2 + 1 = 0$, so its eigenvalues are $\pm i$, which are neither $+1$ nor -1. In general a real orthogonal matrix has eigenvalues of modulus 1, but they need not be real.
6. (D) is true. $\det(A)^2 = \det(A^T)\det(A) = \det(A^TA) = \det(I) = 1$, so $\det A = \pm 1 \neq 0$. Thus A is invertible and $\mathrm{rank}(A) = n$, i.e. full rank.
7. Correct options: (A) and (D).

KEY IDEA
$\|Ax\|^2 = \|x\|^2$ for all x forces $A^TA = I$ (a symmetric matrix whose quadratic form vanishes for every x is zero). So A is orthogonal, hence invertible with $\det A = \pm 1$. Eigenvalues of an orthogonal matrix have modulus 1 but can be complex.

WATCH OUT FOR
Choosing (C) because orthogonal matrices have eigenvalues of absolute value 1. Over the reals they can be complex, as for a rotation with eigenvalues $\pm i$. Another slip is choosing (B) by reading the property as 'A leaves every vector unchanged'. It only says that lengths are preserved.

Q.53 MSQ · 2 MARKS

Machine Learning › LDA, SVM & trees

Consider designing a linear binary classifier $f(x) = \mathrm{sign}(w^\top x + b)$, $x \in \mathbb{R}^2$ on the following training data: Class-1: $\{(2,0)^\top, (0,2)^\top, (2,2)^\top\}$, Class-2: $\{(0,0)^\top\}$. Hard-margin support vector machine (SVM) formulation is solved to obtain w and b. Which of the following options is/are correct?

(A) $w = (4,4)^\top$ and $b = 1$

(B) The number of support vectors is 3

(C) The margin is $\sqrt{2}$

(D) Training accuracy is 98%

Official answer: **B, C**

SOLUTION

1. Label Class-1 as $y = +1$ and Class-2 as $y = -1$. The hard-margin SVM is: minimise $\frac{1}{2}\|w\|^2$ subject to $y_i(w^\top x_i + b) \geq 1$ for all four points.
2. The four constraints are: $(0,0)$ gives $-b \geq 1$, so $b \leq -1$. $(2,0)$ gives $2w_1 + b \geq 1$. $(0,2)$ gives $2w_2 + b \geq 1$. $(2,2)$ gives $2w_1 + 2w_2 + b \geq 1$.
3. Since $b \leq -1$, the second constraint gives $w_1 \geq \dfrac{1-b}{2} \geq 1$, and the third gives $w_2 \geq \dfrac{1-b}{2} \geq 1$. Hence for every feasible (w, b) we have $\|w\|^2 = w_1^2 + w_2^2 \geq 2$.
4. The bound is attained at $w = (1,1)^\top$, $b = -1$. Check all constraints: $(0,0)$: $-(0-1) = 1 \geq 1$; $(2,0)$: $2 - 1 = 1 \geq 1$; $(0,2)$: $2 - 1 = 1 \geq 1$; $(2,2)$: $4 - 1 = 3 \geq 1$. So this is the optimum, and it is unique because the minimum of $w_1^2 + w_2^2$ over $w_1, w_2 \geq 1$ occurs only at $(1,1)$.
5. Support vectors are the points whose constraint is tight, $y_i(w^\top x_i + b) = 1$. These are $(0,0)$, $(2,0)$ and $(0,2)$; the point $(2,2)$ has value $3 > 1$, so it is not one. The number of support vectors is 3, so (B) is correct.
6. Margin width $= \dfrac{2}{\|w\|} = \dfrac{2}{\sqrt{2}} = \sqrt{2}$. Check: the line $x_1 + x_2 = 1$ is at distance $\frac{1}{\sqrt{2}}$ from each of the two supporting lines $x_1 + x_2 = 0$ and $x_1 + x_2 = 2$, so the gap is $\sqrt{2}$. So (C) is correct.
7. (A) is false: with $w = (4,4)^\top$, $b = 1$ the point $(0,0)$ gives $y(w^\top x + b) = -1 \not\geq 1$, so it is not even feasible (and it is not the optimum). (D) is false: the data are linearly separable, so the hard-margin SVM classifies all 4 points correctly, giving training accuracy 100%, not 98% (with 4 points, accuracy can only be 0, 25, 50, 75 or 100 percent).

KEY IDEA

Write the four margin constraints. The point $(0,0)$ forces $b \leq -1$, which forces $w_1, w_2 \geq 1$, so the smallest norm is at $w = (1,1)$ with $b = -1$. Support vectors are the points whose constraint holds with equality, and the margin width is $2/\|w\|$.

WATCH OUT FOR

Counting $(2,2)$ as a support vector because it is in Class-1, or taking the margin as $1/\|w\| = 1/\sqrt{2}$ instead of the full width $2/\|w\|$. Another slip is trusting option (A) without checking feasibility at $(0,0)$.

Q.54 MSQ · 2 MARKS

Probability and Statistics › Random variables

Consider a coin-toss experiment where the probability of head showing up is p. In the i^{th} coin toss, let $X_i = 1$ if head appears, and $X_i = 0$ if tail appears. Consider $\hat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i$ where n is the total number of independent coin tosses. Which of the following statements is/are correct?

(A) $E[\hat{p}] = p$

(B) $E[\hat{p}] = \frac{p}{n}$

(C) As n increases, variance of $\hat{p}$ decreases

(D) Variance of $\hat{p}$ does not depend on n

Official answer: **A, C**

SOLUTION

1. Each X_i is a Bernoulli(p) variable, so $E[X_i] = p$ and $\mathrm{Var}(X_i) = p(1-p)$.

2. Mean, by linearity of expectation: $E[\hat{p}] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \frac{1}{n} \cdot np = p$. So $\hat{p}$ is unbiased: (A) is true and (B) is false, because $\frac{p}{n}$ would only equal p when $n = 1$ (or $p = 0$).
3. Variance: the tosses are independent, so variances add and $\mathrm{Var}(cY) = c^2 \mathrm{Var}(Y)$. Hence $\mathrm{Var}(\hat{p}) = \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}(X_i) = \frac{1}{n^2} \cdot n\,p(1-p) = \frac{p(1-p)}{n}$.
4. The variance has n in the denominator, so for fixed p with $0 < p < 1$ it decreases as n increases: (C) is true, and (D) is false because the variance depends on n.
5. Correct options: (A) and (C).

KEY IDEA

The sample proportion is the mean of n independent Bernoulli(p) variables. Its mean is p (unbiased), and its variance is $\frac{p(1-p)}{n}$, which shrinks as n grows.

WATCH OUT FOR

Confusing the mean of the average with the mean of the sum. $E[\sum X_i] = np$, but dividing by n gives p, not $\frac{p}{n}$. Another slip is forgetting that $\frac{1}{n}$ is squared in the variance, which gives $\frac{1}{n^2} \cdot n = \frac{1}{n}$ instead of $\frac{1}{n} \cdot n = 1$.

Q.55 MSQ · 2 MARKS

Machine Learning › LDA, SVM & trees

Consider a two-class problem in $\mathbb{R}^d$ with class labels red and green. Let μ_{red} and μ_{green} be the means of the two classes. Given test sample $x \in \mathbb{R}^d$, a classifier calculates the squared Euclidean distance (denoted by $\|\cdot\|^2$) between x and the means of the two classes and assigns the class label that the sample x is closest to. That is, the classifier computes $f(x) = \|\mu_{red} - x\|^2 - \|\mu_{green} - x\|^2$ and assigns the label red to x if $f(x) < 0$, and green otherwise. Which of the following statements is/are correct?

(A) The sample $x = 0$ is assigned the label green if $\|\mu_{red}\| < \|\mu_{green}\|$

(B) f is a linear function of x

(C) $f(x) = w^\top x + b$, where w and b are functions of μ_{red} and μ_{green}

(D) f is a quadratic polynomial in x

Official answer: **B, C**

SOLUTION

1. Expand each squared distance: $\|\mu - x\|^2 = \|\mu\|^2 - 2\mu^\top x + \|x\|^2$.
2. Subtract the two expansions. The $\|x\|^2$ terms cancel, so
$$f(x) = \|\mu_{red}\|^2 - \|\mu_{green}\|^2 - 2\mu_{red}^\top x + 2\mu_{green}^\top x.$$
3. Hence $f(x) = w^\top x + b$ with $w = 2(\mu_{green} - \mu_{red})$ and $b = \|\mu_{red}\|^2 - \|\mu_{green}\|^2$. Both w and b depend only on the two class means, so f is linear (affine) in x: (B) and (C) are correct.
4. (D): no term of degree 2 in x survives the subtraction, so f has degree 1 and is not a quadratic polynomial. (D) is false.
5. (A): at $x = 0$, $f(0) = b = \|\mu_{red}\|^2 - \|\mu_{green}\|^2$. If $\|\mu_{red}\| < \|\mu_{green}\|$ then $f(0) < 0$, so the label is red, not green. (A) is false.
6. Correct options: B and C.

KEY IDEA

The nearest-mean classifier with squared Euclidean distance has a linear decision rule: the $\|x\|^2$ terms cancel in the difference, leaving a hyperplane boundary $w^\top x + b = 0$.

WATCH OUT FOR

Seeing squared distances and concluding f is quadratic in x. The quadratic term $\|x\|^2$ is identical in both distances and cancels. Another slip is reading $f(0) < 0$ as green, when negative f means red.

Q.56 MSQ · 2 MARKS

Database Management and Warehousing › Normal forms

Consider the following two relations, named `Customer` and `Person`, in a database:

```
Person (
  aadhaar CHAR(12) PRIMARY KEY,
  name VARCHAR(32));
Customer (
  name VARCHAR(32),
  email VARCHAR(32) PRIMARY KEY,
  phone CHAR(10),
  aadhaar CHAR(12),
  FOREIGN KEY (aadhaar) REFERENCES Person(aadhaar));
```

Which of the following statements is/are correct?

(A) `aadhaar` is a candidate key in the `Customer` relation

(B) `phone` can be NULL in the `Customer` relation

(C) `aadhaar` is a candidate key in the `Person` relation

(D) `aadhaar` can be NULL in the `Person` relation

Official answer: **B, C**

SOLUTION

1. A candidate key is a minimal set of attributes that uniquely identifies every tuple of a relation. A `PRIMARY KEY` is always a candidate key, and it is implicitly `UNIQUE` and `NOT NULL`.
2. `Person`: `aadhaar` is declared `PRIMARY KEY`, so its values are unique and never NULL. It is a candidate key of `Person`, so (C) is correct.
3. Because `aadhaar` is the primary key of `Person`, it can never be NULL there, so (D) is false.
4. `Customer`: the primary key is `email`. The column `aadhaar` is only a foreign key, and a foreign key carries no uniqueness constraint. Several customers can share one `aadhaar` value (for example, one person registered with two e-mail addresses), so `aadhaar` does not identify tuples uniquely. (A) is false.
5. `Customer.phone` is declared as `CHAR(10)` with no `NOT NULL` and is not part of the primary key. Only primary-key columns are forced to be non-NULL, so `phone` can be NULL. (B) is correct.
6. Correct options: (B) and (C).

KEY IDEA

A primary key is a candidate key and is automatically unique and NOT NULL. A foreign key column gives neither uniqueness nor non-nullness, so it is a candidate key only if something else makes it unique.

WATCH OUT FOR

Treating the foreign key `Customer.aadhaar` as a candidate key because it refers to a primary key. The reference makes values in `Customer.aadhaar` match values in `Person`, but it does not stop two customers from sharing a value.

Q.57 MSQ · 2 MARKS

Database Management and Warehousing › Normal forms

Consider a database relation R with attributes ABCDEFG, and having the following functional dependencies:

A → BCEF E → DG BC → A

Which of the following statements is/are correct?

(A) A is the only candidate key of R

(B) A, BC are the candidate keys of R

(C) A, BC, E are the candidate keys of R

(D) Relation R is not in Boyce-Codd Normal Form (BCNF)

Official answer: **B, D**

SOLUTION

1. Closure of A: $A \to BCEF$ gives $\{A, B, C, E, F\}$, then $E \to DG$ adds D, G. So $A^+ = ABCDEFG$, which is all of R. A is a superkey, and being a single attribute it is a candidate key.
2. Closure of BC: $BC \to A$ gives A, and A^+ is all of R, so $(BC)^+ = ABCDEFG$. Also $B^+ = \{B\}$ and $C^+ = \{C\}$ since no FD has B or C alone on its left side. So BC is a minimal superkey, hence a candidate key. Therefore BC is also a candidate key, and (A) is false.
3. Closure of E: $E \to DG$ gives $E^+ = \{E, D, G\}$, which is not all of R. So E is not a key, and (C) is false.
4. Are there any other candidate keys? The only FDs that put A on the right side are $BC \to A$. So a set that does not contain A can reach A only if it contains both B and C. Such a set contains BC and is not minimal. Hence the candidate keys are exactly $\{A\}$ and $\{B, C\}$, and (B) is true.
5. BCNF needs the left side of every non-trivial FD to be a superkey. $A \to BCEF$ and $BC \to A$ satisfy this. $E \to DG$ is non-trivial and E is not a superkey (step 3), so it violates BCNF.
6. So R is not in BCNF, and (D) is true. The correct options are (B) and (D).

KEY IDEA

Find candidate keys by computing attribute closures. Then check every non-trivial FD for a superkey on its left side to test BCNF. Any key not containing A must contain BC, since BC → A is the only way to derive A.

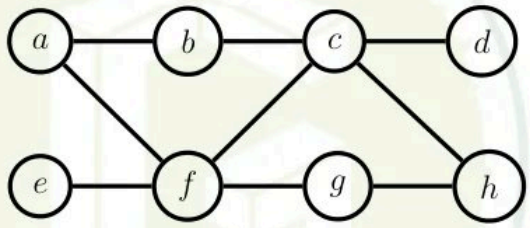
WATCH OUT FOR

Stopping at the key A and calling it the only candidate key, which misses that BC → A makes BC a key too. Another slip is treating E as a key because it determines D and G, when E+ = {E, D, G} does not include A, B, C or F.

Q.58 MSQ · 2 MARKS

Data Structures and Algorithms › Graph algorithms

Q. 58 Let G be a simple, unweighted, and undirected graph. A subset of the vertices and edges of G are shown below.



It is given that $a-b-c-d$ is a shortest path between a and d; $e-f-g-h$ is a shortest path between e and h; $a-f-c-h$ is a shortest path between a and h. Which of the following is/are NOT the edges of G?

(A) (b, d)

(B) (b, g)

(C) (b, h)

(D) (e, g)

Official answer: **A, C, D**

SOLUTION

1. The figure shows the edges $a-b$, $b-c$, $c-d$, $e-f$, $f-g$, $g-h$, $a-f$, $c-f$ and $c-h$. In an unweighted graph the length of a path is its number of edges, and a path is a shortest path only if no edge of G gives a shorter route between its end points.
2. $a-b-c-d$ is a shortest path, so $d(a, d)=3$. If (b, d) were an edge, $a-b-d$ would have length $2<3$. So (b, d) is NOT an edge of G.
3. $e-f-g-h$ is a shortest path, so $d(e, h)=3$. If (e, g) were an edge, $e-g-h$ would have length $2<3$. So (e, g) is NOT an edge of G.
4. $a-f-c-h$ is a shortest path, so $d(a, h)=3$. If (b, h) were an edge, $a-b-h$ would have length $2<3$. So (b, h) is NOT an edge of G.
5. Now test (b, g). Suppose it is added to the drawn edges. The route $a-b-g-h$ has length 3, which equals $d(a, h)=3$, so $a-f-c-h$ is still shortest. Any route from e to h that uses (b, g) must first reach b, which takes at least 3 edges from e ($e-f-a-b$ or $e-f-c-b$), so it is longer than 3. Any route from a to d that uses (b, g) must pass through g and then reach d only via c, so it is longer than 3. All three given shortest-path conditions still hold, so (b, g) is allowed to be an edge. It is not forced to be absent.
6. Hence the pairs that are definitely NOT edges of G are (b, d), (b, h) and (e, g), which are options (A), (C) and (D).

KEY IDEA
In an unweighted graph a path is shortest only if no edge of the graph creates a shorter route between its end points. Each given shortest path therefore forbids every chord that would shorten it.

WATCH OUT FOR
Marking (b, g) as a non-edge because it looks like a chord. It creates the route $a - b - g - h$ of length 3, which only ties $d(a, h) = 3$ and does not shorten any of the three given shortest paths, so it is allowed. Also forgetting that $d(a, h) = 3$ comes from the third path, which is what rules out (b, h).

Q.59 NAT · 2 MARKS

Calculus and Optimization › Limits & continuity

Let $f : \mathbb{R} \to \mathbb{R}$ be such that $|f(x) - f(y)| \leq (x - y)^2$ for all $x, y \in \mathbb{R}$. Then $f(1) - f(0) =$ _____ (Answer in integer)

Official answer: **0**

SOLUTION

1. Fix x and put $y = x + h$ with $h \neq 0$. The condition gives $|f(x + h) - f(x)| \leq h^2$.
2. Divide by $|h|$: $\left|\dfrac{f(x + h) - f(x)}{h}\right| \leq |h|$.
3. Let $h \to 0$. The right side tends to 0, so by the squeeze theorem the difference quotient tends to 0. Hence $f'(x)$ exists and $f'(x) = 0$ for every $x \in \mathbb{R}$.
4. A function with zero derivative everywhere on $\mathbb{R}$ is constant (mean value theorem: $f(1) - f(0) = f'(c) \cdot 1 = 0$ for some $c \in (0, 1)$). So $f(1) = f(0)$.
5. Check without derivatives: split $[0, 1]$ into n equal steps of length $\dfrac{1}{n}$. The triangle inequality gives $|f(1) - f(0)| \leq \sum_{k=1}^{n} \left|f\left(\frac{k}{n}\right) - f\left(\frac{k-1}{n}\right)\right| \leq n \cdot \dfrac{1}{n^2} = \dfrac{1}{n}$. This holds for every n, and $\dfrac{1}{n} \to 0$, so $|f(1) - f(0)| = 0$.
6. Hence $f(1) - f(0) = 0$.

KEY IDEA
A bound of the form $|f(x) - f(y)| \leq C|x - y|^p$ with $p > 1$ forces f to be constant, because the difference quotient is squeezed to 0 and so the derivative is 0 everywhere.

WATCH OUT FOR
Treating f as an arbitrary function and guessing a value such as 1 from the right-hand side $(1 - 0)^2$. That bound only says $|f(1) - f(0)| \leq 1$. It becomes 0 only after splitting the interval into many small steps, or by showing the derivative is 0.

Q.60 NAT · 2 MARKS

Machine Learning › PCA

Let $D = \{x^{(1)}, \ldots, x^{(n)}\}$ be a dataset of n observations where each $x^{(i)} \in \mathbb{R}^{100}$. It is given that $\sum_{i=1}^{n} x^{(i)} = 0$. The covariance matrix computed from D has eigenvalues $\lambda_i = 100^{2-i}$, $1 \leq i \leq 100$. Let $u \in \mathbb{R}^{100}$ be the direction of maximum variance with $u^\top u = 1$.

The value of $\dfrac{1}{n} \sum_{i=1}^{n} \left(u^\top x^{(i)}\right)^2 =$ _____

(Answer in integer)

Official answer: **100**

SOLUTION

1. The data are centred, $\sum_{i=1}^{n} x^{(i)} = 0$, so the sample mean is $\bar{x} = 0$. The covariance matrix (with the $\frac{1}{n}$ convention, which is the one the quantity in the question matches) is

$$C = \frac{1}{n}\sum_{i=1}^{n} x^{(i)}x^{(i)\top}.$$

2. Expand the required quantity:
$$\frac{1}{n}\sum_{i=1}^{n}\left(u^\top x^{(i)}\right)^2 = \frac{1}{n}\sum_{i=1}^{n} u^\top x^{(i)}x^{(i)\top}u = u^\top\left(\frac{1}{n}\sum_{i=1}^{n} x^{(i)}x^{(i)\top}\right)u = u^\top C u.$$
3. Because the mean is zero, $u^\top C u$ is exactly the variance of the data projected on the unit vector u.
4. C is symmetric, so write $C = \sum_j \lambda_j v_j v_j^\top$ with orthonormal eigenvectors v_j. For a unit vector $u = \sum_j a_j v_j$ with $\sum_j a_j^2 = 1$, we get $u^\top C u = \sum_j \lambda_j a_j^2 \le \lambda_{\max}\sum_j a_j^2 = \lambda_{\max}$. Equality holds when u is the eigenvector of the largest eigenvalue. So the direction of maximum variance gives $u^\top C u = \lambda_{\max}$.
5. The eigenvalues are $\lambda_i = 100^{2-i}$, which strictly decrease as i increases: $\lambda_1 = 100^1 = 100$, $\lambda_2 = 100^0 = 1$, $\lambda_3 = 100^{-1} = 0.01$, and so on. Hence $\lambda_{\max} = \lambda_1 = 100$.
6. Therefore
$$\frac{1}{n}\sum_{i=1}^{n}\left(u^\top x^{(i)}\right)^2 = \lambda_1 = 100.$$

KEY IDEA
For centred data, the mean squared projection on a unit vector u equals $u^\top C u$, the variance along u. The direction of maximum variance (the first principal component) gives the largest eigenvalue of C.

WATCH OUT FOR
Reading $\lambda_i = 100^{2-i}$ with the wrong index and taking $\lambda_1 = 100^2 = 10^4$ or λ_{100} instead of $\lambda_1 = 100^1 = 100$. Another slip is to report the sum of all the eigenvalues, which is the total variance, not the variance along the single top direction.

Q.61 NAT · 2 MARKS

Probability and Statistics › Discrete distributions

A bag contains 5 white balls and 10 black balls. In a random experiment, n balls are drawn from the bag one at a time with replacement. Let S_n denote the total number of black balls drawn in the experiment.

The expectation of S_{100} denoted by $E[S_{100}] =$ _____

(Round off to one decimal place)

Official answer: **66.6 to 66.7**

SOLUTION

1. The bag holds $5 + 10 = 15$ balls, of which 10 are black.
2. Each draw is made with replacement, so the draws are independent and the bag is the same every time. On every draw $P(\text{black}) = \frac{10}{15} = \frac{2}{3}$.
3. Let X_i be the indicator that draw i is black. Then $S_{100} = \sum_{i=1}^{100} X_i$, and $S_{100} \sim \text{Binomial}\left(100, \frac{2}{3}\right)$.
4. By linearity of expectation, $E[S_{100}] = \sum_{i=1}^{100} E[X_i] = np$, because $E[X_i] = P(X_i = 1) = \frac{2}{3}$.
5. $$E[S_{100}] = 100 \times \frac{2}{3} = \frac{200}{3} \approx 66.667.$$
6. Rounded to one decimal place, the answer is 66.7.

KEY IDEA
With replacement, every draw is an independent trial with the same success probability $p = 10/15 = 2/3$. So S_n is Binomial(n, p) and $E[S_n] = np$. Linearity of expectation gives this directly, without needing independence.

WATCH OUT FOR
Using $p = 1/2$, or counting black balls out of 10 instead of out of 15 (total balls). Another slip is treating the draws as without replacement, which does not change the expectation but is a wrong model for this question.

Q.62 NAT · 2 MARKS
Database Management and Warehousing › Relational algebra

Q. 62 Consider the following tables, **Loan** and **Borrower**, of a bank.

Loan		
loan_num	**branch_name**	**amount**
L11	Banjara Hills	90000
L14	Kondapur	50000
L15	SR Nagar	40000
L22	SR Nagar	25000
L23	Balanagar	80000
L25	Kondapur	70000
L19	SR Nagar	65000

Borrower	
customer_name	**loan_num**
Anand	L11
Karteek	L11
Karteek	L14
Ankita	L15
Gopal	L19
Karteek	L22
Karteek	L23
Sunil	L23
Sunil	L25

Query: $\pi_{\textbf{branch_name, customer_name}}(\textbf{Loan} \bowtie \textbf{Borrower}) \div \pi_{\textbf{branch_name}}(\textbf{Loan})$

where $\bowtie$ denotes natural join.

The number of tuples returned by the above relational algebra query is _____

(Answer in integer)

Official answer: **1**

SOLUTION

1. Division $R \div S$ returns every value of the attribute that is in R but not in S (here customer_name) that is paired in R with **every** tuple of S. The query is $\pi_{\text{branch_name, customer_name}}(\text{Loan} \bowtie \text{Borrower}) \div \pi_{\text{branch_name}}(\text{Loan})$.
2. The divisor is $S = \pi_{\text{branch_name}}(\text{Loan}) = \{\text{Banjara Hills}, \text{Kondapur}, \text{SR Nagar}, \text{Balanagar}\}$. Duplicates collapse, so S has 4 branches.
3. Natural join on loan_num , row by row: L11 gives (Banjara Hills, Anand) and (Banjara Hills, Karteek); L14 gives (Kondapur, Karteek); L15 gives (SR Nagar, Ankita); L19 gives (SR Nagar, Gopal); L22 gives (SR Nagar, Karteek); L23 gives (Balanagar, Karteek) and (Balanagar, Sunil); L25 gives (Kondapur, Sunil). After projection onto (branch, customer) there are 9 distinct tuples.
4. Group the branches by customer: Anand $\{\text{Banjara Hills}\}$; Ankita $\{\text{SR Nagar}\}$; Gopal $\{\text{SR Nagar}\}$; Sunil $\{\text{Balanagar}, \text{Kondapur}\}$; Karteek $\{\text{Banjara Hills}, \text{Kondapur}, \text{SR Nagar}, \text{Balanagar}\}$.
5. A customer is in the result only if their set of branches contains all 4 branches of S. Karteek does. Sunil lacks Banjara Hills and SR Nagar, and the others each have one branch.
6. The result is $\{\text{Karteek}\}$, so the query returns 1 tuple.

KEY IDEA

Division keeps the customers whose set of branches (after the join) contains every branch of the divisor. Build the joined (branch, customer) pairs, group them by customer, and compare each group with the full branch set.

WATCH OUT FOR

Taking the divisor to be the branches of one particular customer, or counting the branch list with duplicates instead of as a set. Another slip is missing that Karteek reaches SR Nagar through L22 and Balanagar through L23, so he really does cover all four branches.

Q.63 NAT · 2 MARKS

Data Structures and Algorithms › Python

Consider the following Python code snippet.

```
def f(a,b):
  if (a==0):
     return b
  if (a%2==1):
     return 2*f((a-1)/2,b)
  return b+f(a-1,b)
print(f(15,10))
```

The value printed by the code snippet is _____ (Answer in integer)

Official answer: **160**

SOLUTION

1. Start with `f(15,10)`. Here `a = 15`, so `a==0` is false. Also `15 % 2 == 1` is true, so the second `if` fires and the call returns `2*f((15-1)/2, 10)`. In Python 3, `/` is true division, so `(15-1)/2` is the float `7.0`. The first argument is a float from here on, but this changes no branch: `7.0 == 0` is false, `7.0 % 2` is `1.0`, and `1.0 == 1` is true.
2. `f(7.0,10)`: not zero and odd, so it returns `2*f(3.0,10)`, since `(7.0-1)/2 = 3.0`.
3. `f(3.0,10)`: not zero and odd, so it returns `2*f(1.0,10)`, since `(3.0-1)/2 = 1.0`.
4. `f(1.0,10)`: not zero and odd, so it returns `2*f(0.0,10)`, since `(1.0-1)/2 = 0.0`.
5. `f(0.0,10)`: `0.0 == 0` is true in Python, so the first `if` fires and it returns `b = 10`.
6. Every argument on the way down (15, 7.0, 3.0, 1.0) was odd, so the last line `return b+f(a-1,b)` is never reached. Only the doubling branch is used, and it is used exactly four times.
7. Unwind: $f(1.0, 10) = 2 \cdot 10 = 20$, $f(3.0, 10) = 2 \cdot 20 = 40$, $f(7.0, 10) = 2 \cdot 40 = 80$, $f(15, 10) = 2 \cdot 80 = 160$. The program prints `160.0`, since the result of `2*f(7.0,10)` is a float. The integer answer is 160, which equals $b \cdot 2^4 = 10 \cdot 16$.

KEY IDEA

Trace the recursion and watch the parity of the first argument. 15, 7, 3 and 1 are all odd, so each call just doubles the result of the next, down to the base case `f(0,b) = b`. The answer is $b \cdot 2^k$, where $k = 4$ is the number of halvings from 15 to 0. Python 3's `/` makes the arguments floats, but `7.0 % 2 == 1` and `0.0 == 0` still behave as for integers.

WATCH OUT FOR

Taking the final line `b+f(a-1,b)` as if it runs at every level, which gives a sum instead of a product. Another slip is to worry that the float arguments (`7.0`, and so on) break the `a%2==1` or `a==0` tests. They do not, because float values compare equal to the matching integers.

Q.64 NAT · 2 MARKS

Data Structures and Algorithms › Linear structures

Consider the following pseudocode.

```
 Create empty stack S
Set x=0, flag=0, sum=0
Push x onto S
while (S is not empty){
  if (flag equals 0){
    Set x = x+1
    Push x onto S}
  if (x equals 8):
    Set flag=1
  if (flag equals 1){
    x = Pop(S)
    if (x is odd):
      Pop(S)
    Set sum = sum + x}
}
Output sum
```

The value of `sum` output by a program executing the above pseudocode is _____ (Answer in integer)

Official answer: **24**

SOLUTION

1. Initially the stack is $S = [0]$ (bottom to top), with $x = 0$, $\text{flag} = 0$, $\text{sum} = 0$.
2. While $\text{flag} = 0$, each pass of the loop executes `x = x+1` and pushes the new x. Nothing is popped. After pass 8, $x = 8$ and $S = [0, 1, 2, \ldots, 8]$. The test `x equals 8` is now true, so `flag=1`.
3. In that same pass 8 the `flag equals 1` block also runs. `x = Pop(S)` gives $x = 8$ and $S = [0, \ldots, 7]$. Since 8 is even there is no second pop, and $\text{sum} = 0 + 8 = 8$.
4. From pass 9 on, $\text{flag} = 1$, so the first `if` never runs again and nothing is pushed. The test `x equals 8` may set `flag=1` again, which changes nothing. Each pass pops once into x, and pops a second element (discarded) if that x is odd. The popped x is the value added to `sum`.
5. Pass 9: pop 7 (odd), so also pop 6. $\text{sum} = 8 + 7 = 15$, $S = [0, \ldots, 5]$.
6. Pass 10: pop 5 (odd), so also pop 4. $\text{sum} = 15 + 5 = 20$, $S = [0, 1, 2, 3]$.
7. Pass 11: pop 3 (odd), so also pop 2. $\text{sum} = 20 + 3 = 23$, $S = [0, 1]$.
8. Pass 12: pop 1 (odd), so also pop 0. $\text{sum} = 23 + 1 = 24$, and S is empty, so the loop ends.
9. Output: $\text{sum} = 8 + 7 + 5 + 3 + 1 = 24$.

KEY IDEA

Split the run into two phases. First, with $\text{flag} = 0$, the stack fills with $0, 1, \ldots, 8$. Then with $\text{flag} = 1$, each pass pops one element and also discards the next one whenever the popped value is odd. Note that the pass in which x reaches 8 also performs the first pop, because the `flag equals 1` test comes after `flag` is set in the same pass.

WATCH OUT FOR

Treating the pass that sets `flag=1` as a pure push pass and so skipping the pop of 8. That leaves 8 out of the sum and gives 16 instead of 24. Another slip is to forget that the second pop after an odd value discards the next element without adding it to the sum.

Q.65 NAT · 2 MARKS

Data Structures and Algorithms › Graph algorithms

Consider a directed graph $G = (V, E)$, where $V = \{0, 1, 2, \ldots, 100\}$ and $E = \{(i, j) : 0 < j - i \leq 2, \text{ for all } i, j \in V\}$. Suppose the adjacency list of each vertex is in decreasing order of vertex number, and

depth-first search (DFS) is performed at vertex 0. The number of vertices that will be discovered after vertex 50 is _____ (Answer in integer)

Official answer: **75**

SOLUTION

1. Each vertex i has edges only to $i+1$ and $i+2$ (those that exist in V). In decreasing order the adjacency list of i is $(i+2,\ i+1)$, so DFS always tries $i+2$ first.
2. DFS from 0 therefore descends the even chain $0 \to 2 \to 4 \to \cdots \to 100$. Vertex 100 has no out-edges, so DFS backtracks.
3. The even vertices discovered after 50 on this descent are $52, 54, \ldots, 100$, which is $\frac{100-52}{2}+1 = 25$ vertices.
4. Back at 98: its list is $(100, 99)$. Vertex 100 is already discovered, and 99 is not, so 99 is discovered. Its neighbours (100, and 101 which does not exist) give nothing new, so DFS returns to 98 and then to 96.
5. At every even vertex $2k$ on the way back, the first neighbour $2k+2$ is already discovered and the second neighbour $2k+1$ is not, so $2k+1$ is discovered. Its neighbours $2k+3$ and $2k+2$ are already discovered, so DFS returns. This happens at $98, 96, \ldots, 0$.
6. Hence the odd vertices $99, 97, \ldots, 1$ are discovered one by one during the backtrack, in that order. All of them are discovered after the descent has passed 50, so all 50 odd vertices come after vertex 50.
7. Vertices discovered after 50: $25 + 50 = 75$.

KEY IDEA

With adjacency lists in decreasing order, DFS first runs straight down the even chain to 100, then discovers each odd vertex while backtracking, so the discovery order is fixed by 'always take the largest neighbour first'.

WATCH OUT FOR

Counting only the even vertices 52 to 100 (25), or forgetting that the odd vertices below 50 (1 to 49) are also discovered after 50, because they are only reached during backtracking, long after the descent passed 50.

FROM IAE ACADEMY

Preparing for GATE DA?

IAE Academy teaches GATE DA in Ballygunge, Kolkata, and online. Counselling is free: ask us about batches, fees or your preparation.

FREE COUNSELLING

Chat on WhatsApp

Call +91 84204 36524

Open every day, 10:30 am to 7:30 pm.

Scan to chat on WhatsApp

THE GATE DA COURSE

Classroom batches in Ballygunge and online batches. Fees and details are on the course page.

Scan: course page

See the course

FREE ON IAEACADEMY.IN

More GATE DA papers, question by question with the official answers, and free notes by subject.

Scan: more papers

More solved papers

Found this useful? Send it to a friend preparing for GATE.

Send to a friend on WhatsApp

IAE Academy free study material. The questions are those of the GATE DA 2025 question paper, organised by IIT Roorkee; the answers are from its official answer key. Prepared and published by IAE Academy, Kolkata. Free to read and free to share, unaltered and with attribution. Not affiliated with, endorsed by or produced by IIT Roorkee or the GATE committee.

IAE ACADEMY INSTITUTE FOR ENGINEERS

GATE coaching in Kolkata since 2012 · EC, EE, IN, CS and IT, DA, BT and CE · in the classroom and online

227A Rash Behari Ave, Ballygunge, Kolkata 700019

+91 84204 36524 · iaeacademy.in · Updated 19 October 2026

Get directions