📝 真题汇总

Quiz #1 + Midterm | 含完整答案

📌 Quiz #1（Oct 14, 2025）

总分：50分 | 时间：40分钟

Question 1（25分）- Gradient Descent

题目：

We train a simple linear regression model: **y = ax + b**

Given data:

x	y
1	3
2	5

Initial parameters: a = 0, b = 0

Learning rate: η = 0.1

Loss function (MSE):

$$E(a,b) = \frac{1}{n}\sum(y_i - \hat{y}_i)^2$$

where $\hat{y}_i = ax_i + b$ and n = 2

Task: Compute one iteration of Gradient Descent

Show: prediction, error, gradients, and updated values of a and b.

✅ Solution:

Step 1: Predictions

$$\hat{y}_1 = a \cdot x_1 + b = 0 \cdot 1 + 0 = 0$$

$$\hat{y}_2 = a \cdot x_2 + b = 0 \cdot 2 + 0 = 0$$

Step 2: Errors

$$e_1 = \hat{y}_1 - y_1 = 0 - 3 = -3$$

$$e_2 = \hat{y}_2 - y_2 = 0 - 5 = -5$$

Step 3: Gradient (∂E/∂a)

$$\frac{\partial E}{\partial a} = \frac{2}{n}\sum(e_i \cdot x_i)$$

$$= \frac{2}{2}[(e_1 \cdot x_1) + (e_2 \cdot x_2)]$$

$$= \frac{2}{2}[(-3 \cdot 1) + (-5 \cdot 2)]$$

$$= \frac{2}{2}[-3 - 10] = \frac{2}{2} \times (-13) = -13$$

Step 4: Gradient (∂E/∂b)

$$\frac{\partial E}{\partial b} = \frac{2}{n}\sum e_i$$

$$= \frac{2}{2}[e_1 + e_2]$$

$$= \frac{2}{2}[(-3) + (-5)]$$

$$= \frac{2}{2} \times (-8) = -8$$

Step 5: Updates

$$a_{new} = a_{old} - \eta \cdot \frac{\partial E}{\partial a}$$

$$= 0 - 0.1 \times (-13) = 0 + 1.3 = 1.3$$

$$b_{new} = b_{old} - \eta \cdot \frac{\partial E}{\partial b}$$

$$= 0 - 0.1 \times (-8) = 0 + 0.8 = 0.8$$

Final Answer:

a_new = 1.3, b_new = 0.8

New model: y = 1.3x + 0.8

Question 2（25分）- Fuzzy Logic

题目：

Fuzzy sets for Temperature:

- Low: triangular (0, 0, 25)
- Medium: triangular (20, 30, 40)
- High: triangular (35, 50, 50)

Fan speed outputs:

- Slow = 20
- Medium = 50
- Fast = 80

Rules:

1. IF Temperature is Low THEN Speed is Slow
2. IF Temperature is Medium THEN Speed is Medium
3. IF Temperature is High THEN Speed is Fast

Input: Temperature = 30°C

Tasks:

(a) Compute the degree of membership of Temperature = 30°C in each fuzzy set (Low, Medium, High). Show your calculations using the triangular membership functions.

(b) Using the centroid (weighted average) method, compute the defuzzified fan speed output.

✅ Solution:

(a) Membership Degrees

Fuzzy Set	Parameters (a, b, c)	μ(T=30)	Calculation
Low	(0, 0, 25)	**0**	30 > 25 (c) → 0

Medium	(20, 30, 40)	**1**	30 = b (peak) → 1
High	(35, 50, 50)	**0**	30 < 35 (a) → 0

(b) Defuzzification (Centroid Method)

Firing Strengths = Membership Degrees:

Rule	Firing Strength	Output
Rule 1: Low → Slow	0	20
Rule 2: Medium → Medium	1	50
Rule 3: High → Fast	0	80

Centroid Formula:

$$\text{Fan Speed} = \frac{\sum(\text{FS}_i \times \text{Output}_i)}{\sum \text{FS}_i}$$

$$= \frac{(0 \times 20) + (1 \times 50) + (0 \times 80)}{0 + 1 + 0}$$

$$= \frac{0 + 50 + 0}{1} = \frac{50}{1} = 50$$

Final Answer:

Fan Speed = 50

📌 Midterm Exam（Oct 14, 2025）

总分：50分（计算题）+ 50分（MCQ）| 时间：50分钟（计算）+ 15分钟（MCQ）

Question 1（25分）- Multivariate Linear Regression + GD + L2

题目：

You are working with a multivariate linear regression model with the hypothesis function:

$$y = \theta_0 + \theta_1 x_1 + \theta_2 x_2$$

Given:

- A single data point: $(x_1, x_2, y) = (1, 3, 8)$
- Current model parameters: $\theta = [\theta_0, \theta_1, \theta_2] = [1, 0.5, 1]$
- Learning rate for Gradient Descent: $\eta = 0.05$
- L2 Regularization parameter: $\lambda = 0.5$

Tasks:

(i) Using the single data point, perform one iteration of Gradient Descent.

- a) Calculate the prediction $\hat{y}$
- b) Calculate the prediction error
- c) Calculate the updated parameter vector θ_new after one GD step

(ii) Using the initial parameters $\theta = [1, 0.5, 1]$, calculate the L2 regularized cost J_Ridge. (Assume we do not regularize θ_0).

✅ Solution:

(i) Gradient Descent

a) Prediction:

$$\hat{y} = \theta_0 + \theta_1 x_1 + \theta_2 x_2$$
$$= 1 + 0.5 \times 1 + 1 \times 3$$
$$= 1 + 0.5 + 3 = 4.5$$

b) Error:

$$e = \hat{y} - y = 4.5 - 8 = -3.5$$

c) Gradients (MSE only):

$$\frac{\partial \text{MSE}}{\partial \theta_0} = \frac{2}{n}\sum e_i = 2 \times (-3.5) = -7$$

$$\frac{\partial \text{MSE}}{\partial \theta_1} = \frac{2}{n}\sum (e_i \times x_{1i}) = 2 \times (-3.5) \times 1 = -7$$

$$\frac{\partial \text{MSE}}{\partial \theta_2} = \frac{2}{n}\sum (e_i \times x_{2i}) = 2 \times (-3.5) \times 3 = -21$$

Updates (without regularization in this part):

$$\theta_{0,new} = 1 - 0.05 \times (-7) = 1 + 0.35 = 1.35$$

$$\theta_{1,new} = 0.5 - 0.05 \times (-7) = 0.5 + 0.35 = 0.85$$

$$\theta_{2,new} = 1 - 0.05 \times (-21) = 1 + 1.05 = 2.05$$

(ii) L2 Regularized Cost

MSE:

$$\text{MSE} = \frac{1}{n}\sum (y_i - \hat{y}_i)^2 = \frac{1}{1}(8 - 4.5)^2 = 3.5^2 = 12.25$$

L2 Penalty (不包括 θ₀):

$$\text{L2 Penalty} = \lambda \sum_{j=1}^{2} \theta_j^2 = 0.5 \times (\theta_1^2 + \theta_2^2)$$

$$= 0.5 \times (0.5^2 + 1^2) = 0.5 \times (0.25 + 1) = 0.5 \times 1.25 = 0.625$$

Total Cost:

$$J_{Ridge} = \text{MSE} + \text{L2 Penalty}$$

$$= 12.25 + 0.625 = 12.875$$

Final Answers:

(i) **θ_new = [1.35, 0.85, 2.05]**

(ii) **$J_{Ridge} = 12.875$**

Question 2（25分）- Fuzzy Logic (Coffee Maker)

题目：

System: Smart Coffee Maker Strength Control

Input Variables:

1. Coffee Bean Freshness (days since roast):

- Fresh: trapmf(0, 0, 3, 5)
- Medium: trapmf(3, 5, 10, 14)
- Old: trapmf(10, 14, 21, 21)

2. Water Quality (ppm minerals):

- Soft: trapmf(0, 0, 50, 100)
- Medium: trapmf(50, 100, 150, 200)
- Hard: trapmf(150, 200, 300, 300)

Output Variable: Brew Strength:

- Mild: trapmf(0, 0, 3, 4)
- Balanced: trapmf(3, 4, 6, 7)
- Strong: trapmf(6, 7, 10, 10)

Rules:

1. IF Beans are Fresh AND Water is Soft THEN Strength is Mild
2. IF Beans are Medium AND Water is Medium THEN Strength is Balanced
3. IF Beans are Old AND Water is Hard THEN Strength is Strong

Current Input:

- Bean Freshness = 6 days
- Water Quality = 120 ppm

Task: Calculate the firing strength of each rule

✅ Solution:

Step 1: Membership Values for Bean Freshness = 6 days

Fresh: trapmf(0, 0, 3, 5)

6 > d (5) → μ_Fresh(6) = **0**

Medium: trapmf(3, 5, 10, 14)

6 在平顶区间 [b, c] = [5, 10] → μ_Medium(6) = **1**

Old: trapmf(10, 14, 21, 21)

6 < a (10) → μ_Old(6) = **0**

Step 2: Membership Values for Water Quality = 120 ppm

Soft: trapmf(0, 0, 50, 100)

120 > d (100) → μ_Soft(120) = **0**

Medium: trapmf(50, 100, 150, 200)

120 在平顶区间 [b, c] = [100, 150] → μ_Medium(120) = **1**

Hard: trapmf(150, 200, 300, 300)

120 < a (150) → μ_Hard(120) = **0**

Step 3: Firing Strengths (使用 MIN 算子)

Rule	Calculation	Firing Strength
Rule 1: Fresh ∧ Soft → Mild	min(0, 0)	**0**
Rule 2: Medium ∧ Medium → Balanced	min(1, 1)	**1**
Rule 3: Old ∧ Hard → Strong	min(0, 0)	**0**

Final Answer:

Rule 1: 0

Rule 2: 1

Rule 3: 0

只有 Rule 2 被激活!

📌 Midterm MCQ（精选）

50分（10题），15分钟

Question 3 - When to use GD vs Normal Equation?

The primary reason to use Gradient Descent for linear regression instead of the Least Squares Solution is when:

a) the model is severely underfitting.

b) the number of features is very large (e.g., >1000), making LSS computationally expensive.

c) the relationship between variables is perfectly linear.

d) you need a 100% accurate model.

✅ **Answer: (b)**

当特征很多时（>1000），Normal Equation 需要计算 (X^T X)^(-1)，矩阵求逆非常慢，Gradient Descent 更高效。

Question 4 - L1 vs L2 Regularization

You are building a linear regression model and suspect that only 5 out of 100 features are truly predictive. Which regularization technique would be most appropriate to help identify these key features?

a) L2 Regularization (Ridge)

b) L1 Regularization (Lasso)

c) ElasticNet with a higher weight on the L2 part

d) No regularization is needed.

✅ **Answer: (b) L1 Regularization (Lasso)**

L1 (Lasso) 会让某些系数变成 0，起到特征选择的作用。L2 (Ridge) 只会让系数变小，不会变成 0。

Question 6 - Normal Equation 矩阵维度

In the multivariate linear regression normal equation θ = (X^T X)^(-1) X^T Y, if the design matrix X has dimensions m × (n+1) (m examples, n features plus intercept), and Y is m × 1, what are the dimensions of the resulting parameter vector θ?

a) m × 1

b) (n+1) × 1

c) m × (n+1)

d) (n+1) × m

✅ **Answer: (b) (n+1) × 1**

θ 的维度 = 特征数 + 1（截距项）= n+1

Question 7 - Overfitting → Underfitting 转换

A model has high error on both training and test data. Increasing model complexity reduces training error to near zero, but test error remains high. This sequence describes the transition from a model suffering primarily from _____ to one suffering primarily from _____.

a) High variance; High bias

b) High bias; High variance

c) Underfitting; Optimal fitting

d) High bias; Low bias

✅ **Answer: (b) High bias; High variance**

开始：Train ↑ + Test ↑ → Underfitting (High Bias)

增加复杂度后：Train ↓ + Test ↑ → Overfitting (High Variance)

🎯 复习建议

睡前做什么

- ✅ 过一遍公式速查表（文档1）
- ✅ 看一遍这些真题的解题步骤
- ✅ 重点记住：GD 6步、Normal Equation 公式、Fuzzy Logic 计算步骤
- ✅ 早点睡觉，保证状态

明天考试时

- ✅ 先读题，分配时间
- ✅ 先做熟悉的题
- ✅ 写清楚步骤（即使答案错了，步骤对也有分）
- ✅ 检查正负号和单位

🌙 睡个好觉，明天加油！💪