

# 考前速查表

CSCI 6751 | 睡前必看 | 公式 + 概念

⚠ 考试前看一遍！所有公式 + 核心概念

## 1 Gradient Descent (梯度下降)

### 核心公式 (必背!)

更新规则:

$$\theta_{new} = \theta_{old} - \eta \nabla J(\theta)$$

其中:

- $\eta$  (eta) = Learning Rate (学习率)
- $\nabla J$  = 梯度 (损失函数的导数)

### 简单线性回归 ( $y = ax + b$ )

损失函数 (MSE):

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

梯度:

$$\frac{\partial J}{\partial a} = \frac{2}{n} \sum_{i=1}^n (\hat{y}_i - y_i) \cdot x_i$$

$$\frac{\partial J}{\partial b} = \frac{2}{n} \sum_{i=1}^n (\hat{y}_i - y_i)$$

更新:

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

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

### 手算步骤 (6步)

1. 预测:  $\hat{y}_i = a \cdot x_i + b$

2. 误差:  $e_i = \hat{y}_i - y_i$
3. 梯度 **a**:  $\partial J / \partial a = (2/n) \sum (e_i \cdot x_i)$
4. 梯度 **b**:  $\partial J / \partial b = (2/n) \sum (e_i)$
5. 更新 **a**:  $a_{new} = a_{old} - \eta \cdot \partial J / \partial a$
6. 更新 **b**:  $b_{new} = b_{old} - \eta \cdot \partial J / \partial b$

## 多变量线性回归 ( $y = \theta_0 + \theta_1 x_1 + \theta_2 x_2$ )

梯度:

$$\frac{\partial J}{\partial \theta_0} = \frac{2}{n} \sum (\hat{y}_i - y_i)$$

$$\frac{\partial J}{\partial \theta_1} = \frac{2}{n} \sum (\hat{y}_i - y_i) \cdot x_{1i}$$

$$\frac{\partial J}{\partial \theta_2} = \frac{2}{n} \sum (\hat{y}_i - y_i) \cdot x_{2i}$$

## 2 L2 Regularization (Ridge)

### 核心公式

L2 Regularized Cost:

$$J_{Ridge} = \frac{1}{n} \sum (\hat{y}_i - y_i)^2 + \lambda \sum_{j=1}^p \theta_j^2$$

注意：通常不正则化  $\theta_0$ 。（截距项）

梯度（带正则化）:

$$\frac{\partial J}{\partial \theta_j} = \frac{2}{n} \sum (\hat{y}_i - y_i) \cdot x_{ji} + 2\lambda \theta_j$$

### $\lambda$ (Lambda) 的影响

$\lambda$ 大小	影响
$\lambda = 0$	无正则化（普通线性回归）
$\lambda$ 小	轻微惩罚，可能 Overfit
$\lambda$ 中	平衡，最优
$\lambda$ 大	强惩罚，可能 Underfit

### 3 Normal Equation

#### 核心公式（必背！）

$$\theta = (X^T X)^{-1} X^T y$$

#### 步骤

1. 构建设计矩阵  $X$ （第一列全是1）
2. 计算  $X^T X$
3. 求逆  $(X^T X)^{-1}$
4. 计算  $X^T y$
5. 矩阵乘法得到  $\theta$

#### 2×2 矩阵求逆（必须记住！）

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

行列式：

$$\det(A) = ad - bc$$

逆矩阵：

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

口诀：对角交换、副对角取负、除以行列式

#### 何时用 Normal Equation vs GD?

方法	何时用	优缺点
Normal Equation	特征 < 1000	 一步求解  需要求逆，慢

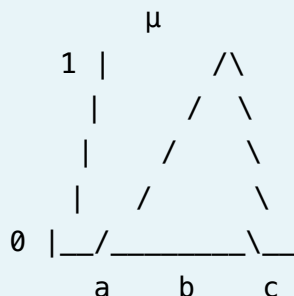
Gradient Descent

特征 > 1000

- ☒ 不需要求逆
- ☐ 需要多次迭代

## 4 Fuzzy Logic (模糊逻辑)

### 三角形隶属函数 $\text{triangular}(a, b, c)$



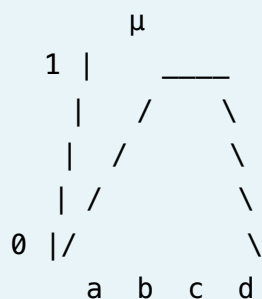
公式:

$$\mu(x) = \begin{cases} 0 & x \leq a \\ \frac{x-a}{b-a} & a < x \leq b \\ \frac{c-x}{c-b} & b < x < c \\ 0 & x \geq c \end{cases}$$

关键点:

- $a$  = 左边界
- $b$  = 峰值 ( $\mu = 1$ )
- $c$  = 右边界

### 梯形隶属函数 $\text{trapezmf}(a, b, c, d)$



公式:

$$\mu(x) = \begin{cases} 0 & x \leq a \\ \frac{x-a}{b-a} & a < x < b \\ 1 & b \leq x \leq c \\ \frac{d-x}{d-c} & c < x < d \\ 0 & x \geq d \end{cases}$$

关键点：

- a 到 b = 上升段
- b 到 c = 平顶 ( $\mu = 1$ )
- c 到 d = 下降段

## Fuzzy Inference System (4步)

1. **Fuzzification** (模糊化)：把清晰输入转为隶属度
2. **Rule Evaluation** (规则评估)：计算 Firing Strength
3. **Aggregation** (聚合)：合并规则输出
4. **Defuzzification** (去模糊化)：转回清晰值

## Firing Strength (激活强度)

IF X is A AND Y is B THEN Z is C

使用 MIN 算子 (AND)：

$$\text{Firing Strength} = \min(\mu_A(x), \mu_B(y))$$

为什么是 MIN? AND = 同时满足，取最弱的

## Centroid Defuzzification (去模糊化)

加权平均法：

$$\text{Output} = \frac{\sum_{i=1}^n (\text{FS}_i \times \text{Output}_i)}{\sum_{i=1}^n \text{FS}_i}$$

其中：

- FS = Firing Strength



- Output = 规则的输出值

## 计算步骤

1. 计算隶属度：对每个输入，计算  $\mu(x)$
2. 计算 **Firing Strength**：  $\min(\mu_1, \mu_2)$  for AND
3. 去模糊化：  $\Sigma(\text{FS} \times \text{Output}) / \Sigma(\text{FS})$

## 5 Overfitting vs Underfitting

### 定义

类型	Train Error	Test Error	原因
Underfitting	高	高	模型太简单
Good Fit	低	低 ( $\approx$ Train)	刚好
Overfitting	很低	高 ( $\gg$ Train)	模型太复杂

### 解决方法

#### 减少 Overfitting:

- ☒ 增加  $\lambda$  (正则化强度)
- ☒ 减少 Polynomial Degree
- ☒ 增加训练数据
- ☒ Early Stopping

#### 减少 Underfitting:

- ☒ 减少  $\lambda$
- ☒ 增加 Polynomial Degree
- ☒ 增加特征

### Hyperparameters (超参数)

- **Learning Rate ( $\eta$ )**: GD 步长
- **Polynomial Degree**: 模型复杂度
- **Lambda ( $\lambda$ )**: 正则化强度

超参数不是从数据学习的, 需要人工调优!

## 6 其他重要概念

### Classification vs Regression

类型	输出	例子
Regression	连续数值	房价、温度
Classification	离散类别	猫狗、垃圾邮件

### Supervised vs Unsupervised Learning

类型	特点	例子
Supervised	有标签	房价预测、猫狗分类
Unsupervised	无标签	客户聚类、降维

### Fuzzy vs Classical Logic

类型	取值	例子
Classical	0 或 1 (二元)	True / False
Fuzzy	0 到 1 (程度)	0.7 (比较热)

## ⚠️ 常见错误（必须避免！）

### Gradient Descent

- ❌ 误差算反了：应该是  $\hat{y} - y$ ，不是  $y - \hat{y}$
- ❌ 忘记除以  $n$ （数据点个数）
- ❌ 更新时加号变减号（应该是减）
- ❌ Learning Rate 忘记乘
- ❌ 梯度对  $a$  的计算忘记乘  $x_i$

### Normal Equation

- ❌ 矩阵大小不匹配就相乘
- ❌ 行列式计算错误（ $ad - bc$ ，不是  $ad + bc$ ）
- ❌ 矩阵求逆时对角元素没交换

### Fuzzy Logic

- ❌ 区间判断错误（上升/下降/平顶/外部）
- ❌ AND 用 MAX（应该用 MIN）
- ❌ Centroid 分子分母算错
- ❌ 梯形平顶区间忘记  $\mu = 1$

## 考试策略

### 时间分配（假设 50 分钟）

- 读题 + 分配时间：3-5 分钟
- Question 1: 20-22 分钟
- Question 2: 20-22 分钟
- 检查：3-5 分钟

### 做题顺序

1. 先做最熟悉的题
2. 再做需要计算的题
3. 最后检查答案

### 答题技巧

- ☒ 写清楚步骤（即使答案错了，步骤对也有分）
- ☒ 检查单位和符号（特别是正负号）
- ☒ 不要在一道题上花太久
- ☒ 用铅笔（方便修改）

🌙 睡前看完这个，明天考试加油！💪

记住：理解 > 记忆，冷静 > 紧张