

面试数学与数据结构累积版

面试概述

总共七类，数学，信号，自控，运筹，电子学，生物信息，数据结构。

2016 年政策：将从至少两类中抽 4 道题，每类至多两道题。可以选择跳过一题。

2019 年政策：先读一段英文并翻译，自己选两类抽 3 道题，可以跳过一题。

2015 及以前

1. 一队人拿着大小水桶排队打水，问：用什么排序方法能够使总的等待时间最少？

答¹：既然不是考排序算法的，那就是小学一年级的题了。从桶最小的开始排队（排序）。当然，如果考虑更高级的情况，给每个桶定义不同的重要性、给每个人的打水速度有额外的规定，那么这道题就比较麻烦了。据说这道题的数学本质是排序不等式。

2. 运筹学：名词解释：计算复杂性

【估计是王书宁出的题，当时他讲了半节课。“首先它分为两种，时间复杂性和空间复杂性....”，然后说了一通 P, NP, NP-hard, NP-complete。】²

答：我不觉得这一定是王书宁出的题，可能是数据结构里的。计算复杂性是一种极限概念，衡量一种算法随着问题规模的增长，对于时间和空间消耗的程度。计算复杂性分为时间复杂性和空间复杂性。对于这两种复杂性的衡量标准为“复杂度”。对于时间复杂度，时间复杂度是时间频度的同等量级函数，用 landau 符号（大O）表达。一般只需要考虑关键语句的执行次数。空间复杂性类似。事实上，这种复杂性的测量也只是一种依赖规模的估计。真正的运行效率还是要依赖运行的物理环境和代码质量。

3. 数据结构：链表，搜索第*i*个节点，增加一个指针，用什么结构可以将效率提高到 $O(\log n)$ 。

【我答得是二叉树，卓晴好像不太满意。】

答：（二叉树有毛用。）这里说的增加指针可能指的是在每个节点结构里增加一个指针。这样的话，应该是二叉排序树，也即二叉搜索树。而且是平衡（或近似平衡）的二叉排序树。完全二叉树、B 树或者 B+ 树或者红黑树都可以。这些树的树高都为 $O(\log n)$ 级别的。普通的二叉搜索树，一旦出现链的情况，那最坏运行时间为 $O(n)$ 。

4. 一个箱子有两个 1 元的，一个 5 元的，问两次取出 6 元的概率是多少，第一次不放回

【我用条件概率算了半天... $\frac{1}{3} + \frac{2}{3} \times \frac{1}{2} = \frac{1}{2}$!!!!】

答：没难度，看当时是否脑残。

5. 说有两个图，节点集合是一样的，图 1 的边是图 2 的边的子集，问这两个图什么关系。

答：学长说“无语了”，其实不太应该啊。如果是在运筹学里抽到这道题，应该能够想到运筹学里“支撑子图”、“支撑树”之类的概念。我觉得数据结构里不太会考图论相关。

6. 一个区间上函数有 0 有非 0，是否保证其一定积分非 0。如果否请说明增加什么条件可以保证。

答：不一定啊，例如

$$f(x) = \begin{cases} 1 & x \in R \setminus Q \\ 0 & x \in Q \end{cases}$$

这个函数的积分就 0，因为有理数集的测度就是 0。应该增加一个连续性条件，一个符号条件，即一个函数在一段区间上是连续的，且有非零值，且在区间上恒为非负或非正。当然这这是一种充分条件。之所以加符号条件，是因为如果有正有负，那就 depends 了。

¹ 为阿狸的回答，下同

² 方括号里为学长们的回忆版回答

7. A 为任意阶矩阵, AA^T 和 A^TA 的秩谁大谁小

答: 显然是一样大的。可以通过 $\text{rank}(A) = \text{rank}(A^T)$ 来证明。可以考虑通过 SVD 分解来证明。

8. 数据结构: 请分析一个算法的空间复杂度和时间复杂度之间的关系

答: (看来对复杂度的概念很重视。) 一个算法的空间复杂度和时间复杂度之间可能有正相关关系, 也可能有负相关关系, 也可能压根没关系 (相互独立), 这取决于具体算法。³

9. 第二个是怎么把三个向量施密特正交化, 记不住公式。要是能记住原理, 可以现推。

答: 只好背公式了。

10. 有 n 个正交向量, 问它们能不能线性相关。

答: (我觉得不存在正交向量这个概念, 只有正交向量组的概念, 因为正交是向量之间的关系。如果说是 n 个向量组成欧式空间中的正交向量组, 那么肯定是线性无关, 见《线性代数与几何(上)》定理 5.20 (p174)。如果仅仅单纯考虑向量的话, 只需要把零向量考虑好就可以了, 零向量与任何向量都正交。

11. 给概率密度函数 $f(x) = 0.5e^{-|x|}$, 求分布函数

答:
$$F(x) = \begin{cases} \frac{1}{2}e^x & x < 0 \\ 1 - \frac{1}{2}e^{-x} & x \geq 0 \end{cases}$$

12. 微积分: 判断一个序列是否绝对收敛, 是否一致收敛。

答: (记住定义多刷题。可能会考比较复杂的表达式, 但应该不会超出书中的范围。后面的题目懒得答了, 后人如果有心可以补充并解答。)

13. 数据结构: 给出一个图, 画出它的广度优先遍历的支撑树 (貌似是支撑树)

答: 运筹学书上有。

14. 给出 a_1, a_2, a_3 , 要求用施密特正交化弄出一组正交基

答: 参考教科书。

15. 数学: 分析可微性, 偏导存在和连续之间的关系

16. 数据结构: 一个 2000 节点的完全二叉树, 问它的深度, 叶子节点数, 有两个子节点的节点数, 只有一个右子节点的节点数。

17. 数据结构: 空间复杂度的定义, 给段代码判断空间复杂度

18. 极坐标变换求一个二维积分 $e^{-(x^2+y^2)}, x \geq 0, y \geq 0$

19. 三个事件相互独立, 充要条件是其中任两个事件彼此独立, 问命题是否正确, 为什么。

20. 矩阵特征值的和与积

21. 判断矩阵相抵的一道选择题, 选项不太记得了, 总之当时看到选项就选了。

22. 时间复杂度的定义, 然后问你时间复杂度就是一段代码的实际执行时间吗, 接下来给一段代码判断时间复杂度。

23. 生成树和最小生成树的定义及区别; 空间复杂度的计算。

24. 问: 什么是完全二叉树, 判断下面两个是不是完全二叉树

25. 数学: 问 A 和 A^TA 的秩的关系, A^TA 和 AA^T 的秩的关系

26. 数据结构: 给了一段代码, 问是什么功能, 很简单, 就是计算叶节点的个数

27. 数学, $|A| < 1, AA^T = I$, 求 $|A + I|$

28. 贝叶斯公式: 甲乙打靶, 甲打中概率为 0.6, 乙打中概率为 0.5, 现知靶被打中, 问是甲打中的概率是多少?

29. X, Y 为两事件, 问 $E(X + Y) = E(X) + E(Y)$ 、 $E(XY) = E(X)E(Y)$ 是否成立;

³ 安亮个人观点。欢迎修正

30. 复数在实数域上是线性空间吗？几维？
31. 已知 A, B 可逆，求 $\begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix}$ 的逆矩阵。
32. 生成树和最小生成树是什么，判断是不是生成树，生成树和最小生成树有哪些应用
33. 二叉排序树是什么，判断是不是二叉排序树
34. 有 3 个结点的树有几种？有 3 个结点的二叉树有几种？
35. A, B 是同阶可逆矩阵。问以下哪个说法正确。我选的是“存在可逆矩阵 P, Q , 使得 $B=PAQ$ ”
36. 数学：两个 1 块钱硬币，三个 2 块钱硬币，问抽取两次，得到三块钱的概率是多少。
37. 数学：对 $\frac{1}{x}$ 积分，积分上下限为 e 和 $-e$ ，然后题目说他等于 $\ln|x|$ ，从 $-e$ 到 e ，然后等于0，问是否正确？
- 答：（我猜的）不对。这个东西是黎曼不可积的。
38. $1, 2, \dots, n$ 号球，摸 k 次，可放回，最后摸出所有球的数字和的数学期望？
- 答：很简单，就是 $\frac{k(n+1)}{2}$
39. 斐波那契数列为什么可以用递归计算？
- 答：（我猜的）因为斐波那契的递推公式里调用了自身。
40. 数据结构：栈、队列、字符串中哪些是线性表，哪些不是，为什么？
- 答：显然都是。因为所有的元素都只有最多一个前驱和一个后继。

2016 年

1. 数据结构：数据结构的四种存储方式是什么？顺序存储和链式存储有什么区别？
- 答：顺序结构，链式结构，索引结构，散列结构。顺序存储的内存空间是连续的，链式存储的内存空间是不连续的。
2. 数据结构：堆的存储结构是什么？逻辑结构是什么？有什么应用？
- 答：顺序结构（数组）。完全二叉树。排序，优先级队列。
3. 数据结构：关系型数据库是什么？数据结构是什么？答案是 B+树。
4. 数据结构：2000 个节点的完全二叉树，问树高，问其中的有两个孩子的节点有多少？
- 答：999
5. 数学： $P(A - B) = P(A) - P(B)$ 的充要条件是什么？
- 答：（我也不确定）根据书上的定理，对于任意的 A 和 B ，有 $P(A - B) = P(A) - P(AB)$ 。所以， $P(AB) = P(B)$ 肯定是充要条件了。那么，这个等式背后的集合含义是 $B \subseteq A$ 。这个应该就是想要的答案了吧。
6. 数学：有 10 个黑球，2 个红球，不放回抽样，则第二次抽到红球的概率是？
- 答： $\frac{1}{6}$
7. 数学：给了一个表达式，问绝对收敛还是条件收敛。
8. 数据结构：栈和队列是不是线性表？
- 答：根据维基百科和其他一些书籍中的定义，线性表只强调前驱和后继的唯一性。所以，栈和队列都是线性表。然而，现场老师说，线性表要求首节点固定，所以队列不是。
9. 数据结构：一个链表，插入节点。给了选项进行排列。
10. 数学：平面积分，算两个圆相交的区域。
11. 数学： L 长度的棍分成两段，问两段长度的相关系数（答案 -1，提醒：别懵逼）
12. 计算二重积分： $\iint_D \sqrt{x^2 + y^2} dx dy$ ，其中 $D: x^2 + y^2 \leq 2y, x^2 + y^2 \geq 2x$ ，选择题，选择正

确的极坐标形式

13. $\int_0^{+\infty} \frac{1}{x^2} e^{\frac{1}{x}} dx$ 判断敛散 (发散)
14. 一个函数在某个区间上可积、非负, 且不恒等于零, 第一问是该函数在该区间上的积分是不是恒大于零, 举例说明; 第二问是什么条件下保证恒大于零 (不是; 大于零的点测度不为零)
15. $\sum_{n=p}^{+\infty} \left(\frac{1}{\sqrt{n}} - \frac{1}{\sqrt{n+1}} \right) \sin(n+k)$ 判敛散
16. A^{-1}, A^T 的特征值与 A 的特征值的关系是什么 (λ^{-1}, λ)
17. AB 互斥 $P(A-B) = P(A)$
18. $\lim_{a \rightarrow 0} \iint_{x^2+y^2 \leq a^2} \frac{1}{\pi a^2} f(x,y) dx dy$ 选择题: A、不存在; B、 $f(0,0)$; C、 $f(2,0)$; D、 $f'(0,0)$
19. $A \in R^{m \times n}, B \in R^{n \times m}$, 选择题, 判断哪种说法正确, 正确的是: $m > n$, 可以推出 $|AB| = 0$
20. $A \subset B$, $P(A)$ 与 $P(A|B)$ 的关系
21. $AB = C$ 可逆, A 与 C 的行秩与列秩的关系
22. $y + xy' = f(x)$ 有三个解 y_1, y_2, y_3 , 问通解是什么?
23. 矩阵特征值和是迹, 秩是行列式
24. 实对称矩阵, 特征向量是否一定正交?
25. $S_N = \sum_{n=1}^N (-1)^n \frac{1}{n^\lambda}$ 收敛性
26. $f(x)$ 可微, $g(x) = f^{-1}(x)$, 求 $\frac{d}{dx} \int_0^{f(x)} tg(t) dt$
27. $f(x,y) = \begin{cases} (x^2+y^2) \sin(\frac{1}{x^2+y^2}) & x^2+y^2 \neq 0 \\ 0 & x^2+y^2 = 0 \end{cases}$, 判断原点的可微性 (课本原题)
28. $X \sim U(0,2)$, $Y = X^2$, 求 $f_Y(y)$
29. $A^2 - 2A - 2I = 0$, 求 $(A - I)^{-1}$

2017 年

1. 循环队列用空出一个元素的方式来区分队满和队空。现在有一个长度为 50 的队列, 问队满时共有 () 元素。若 $front=35, rear=15$, 共有 () 元素。若 $front=15, rear=25$, 共有 () 元素。
- 答案: 49, 31, 11 解析: 其题目意思是, 队满时 $front-rear=49$, 队空时 $front-rear=0$; 其他和普通循环队列没有区别。
2. A 和 B 为任意两事件, 下列选项正确的是
- A. $P(AB) \leq P(A)P(B)$
- B. $P(AB) \geq P(A)P(B)$
- C. $P(AB) \leq \frac{P(A)+P(B)}{2}$
- D. $P(AB) \geq \frac{P(A)+P(B)}{2}$

答案: C 解析: 利用 $P(AB) \leq \min\{P(A), P(B)\}$ 即可求出。另外我一开始问老师是单选还是多选, 老师的回答是你自己判断, 所以所有选项最好都判断一遍, 以防被坑。

3. 一个 200 个节点的二叉树, 最多有多少叶节点? 最小深度是多少?

答案: 100; 8 解析: 最多叶节点和最小深度都是完全二叉树满足的结果, 所以就是求一个

200 节点的完全二叉树，剩下就很好算了。

4. $E(xy) = E(x)E(y)$ 能得到什么结论
5. 给了 4 个数组问那个发散
6. 曲线拟合中的最小二乘
7. N 元非齐次线性方程 $Ax = b$ 的解的个数，和 $Ax = 0$ 的解的个数的关系；（注意 A 未必是方阵）
 $Ax = b$ 无穷多解，则 $Ax = 0$ 无穷多解
 $Ax = b$ 唯一解，则 $Ax = 0$ 唯一解
 $Ax = 0$ 无穷多解，则 $Ax = b$ 无解或无穷多解
 $Ax = 0$ 唯一解，则 $Ax = b$ 无解或唯一解
8. 多元函数的可微，连续，可偏导，偏导连续之间的关系；最好能举出例子；
9. 相关系数是什么；相关系数是 0 能否证明独立性；
10. 相似矩阵的性质判断
11. 无向图中，所有点的度的和与边数的关系；
12. 在 N 个节点的无向图中，至少有多少边才构成连通图；
13. 在 N 个节点的简单有向图中，最多有多少个边；
14. 二叉树的先序遍历，中序遍历，后序遍历
15. 关系型数据库的结构，用的数据结构（B+树），是不是线性表
16. 向量 α, β, γ 线性相关， α, β, δ 线性无关。给出了四个选项，哪些可以被线性表示或不能被线性表示，选择不正确的一个选项。
17. 给出一个函数项级数，描述逐点收敛和一致收敛的概念。
18. 100 次独立重复实验，每次成功概率为 p ，问当 $p = ?$ 时，成功次数的方差最大，这个最大方差是多少？
19. L 为带头表节点的链表， p 不为首元素，请排序以下代码删除 p 的直接前驱：
(a) $p \rightarrow next = p \rightarrow next \rightarrow next;$
(b) $While(p \rightarrow next \rightarrow next != q) p = p \rightarrow next;$
(c) $Q = p;$
(d) $Free(Q)$
(e) $Q = p \rightarrow next;$
(f) $p = L;$

答案：c,f,b,e,a,d

20. 顺序结构是只适用于线性表的存储结构

答案：不对，比如优先级队列采用堆实现，其中堆通过数组实现，为顺序结构，但是逻辑结构为完全二叉树。

21. 若 $E(XY) = E(X)E(Y)$ ，那么下列正确的是：

- A. $D(XY) = D(X)D(Y)$
- B. $D(X - Y) = D(X) - D(Y)$
- C. $X、Y$ 相互独立
- D. $X、Y$ 不相互独立

答案：B

22. 函数如下：

$$f(x, y) = \begin{cases} 1 & xy \neq 0 \\ 0 & xy = 0 \end{cases} \text{问连续 偏导 可微}$$

23. 100 节点 最多叶子结点 最深数树

24. 已知 A, B 独立, 已知 $P(B)$, $P(A - B)$, 求 $P(B - A)$;
25. 微积分求极限, 洛必达;
26. 原题: $n + 1$ 个球抽 k 次, 求编号期望;
27. 原题:三个 1 元两个 2 元, 有放回地抽两次, 求和为 3 元的概率;
28. 哈夫曼 (我忘了定义 跳了)
29. 算法应该从哪些角度评测
30. 已知 A 与 B 相似, 以下不相似的是
 A. A^{-1}, B^{-1}
 B. A^T, B^T
 C. $A^T + A, B^T + B$
 D. $A^{-1} + A, B^{-1} + B$
31. 一根木棒长 L , 拆成两端, 相关系数-1 (往年题)
32. 3 个节点的树有 2 个, 二叉树有 5 个 (往年题)
33. 求解 \int_x^1 在 $[e, -e]$ 区间的积分, 直接使用 $\ln|x| \big|_{-e}^e = 0$, 是否正确, 为什么?
34. 栈、队列, 哪些是线性表, 哪些不是, 为什么
35. $A^2 - 2A - I = 0$, 求 $(A - I)^{-1}$
36. 拓扑排序, AOV 图是什么; 给一个拓扑图判断两个序列是否为正确地拓扑排序序列
37. 一段代码问什么功能复杂度怎样
38. $Ax = b, A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & a \\ 1 & 4 & a^2 \end{bmatrix}, b = \begin{bmatrix} 1 \\ d \\ d^2 \end{bmatrix}, \Omega = \{1, 2\}$ 问 $Ax = b$ 有无穷多解的充要条件是 ()
 A. $a \in \Omega, b \in \Omega$
 B. $a \in \Omega, b \notin \Omega$
 C. $a \notin \Omega, b \in \Omega$
 D. $a \notin \Omega, b \notin \Omega$

2018 年 (无资料)

2019 年

1. 已知 A 为 n 阶矩阵, α 为 n 维向量, 满足

$$B = \begin{bmatrix} A & \alpha \\ \alpha^T & 0 \end{bmatrix}$$

$$r(B) = r(A)$$

选择正确的是

- A. $Ax = \alpha$ 必有唯一解
 B. $Ax = \alpha$ 必有无穷多解
 C. $Bx = 0$ 必有唯一解
 D. $Bx = 0$ 必有无穷多解
2. 若 $E(XY) = E(X)E(Y)$, 那么下列正确的是:
 A. $D(XY) = D(X)D(Y)$
 B. $D(X + Y) = D(X) + D(Y)$
 C. X, Y 相互独立
 D. X, Y 不相互独立
3. 请描述空间复杂度的概念, 空间复杂度的是否指算法实现的程序执行所占空间? 并求矩

阵乘法（朴素实现，时间复杂度 $O(n^3)$ ）算法的空间复杂度。

4. 2000 个节点完全二叉树的深度，叶节点个数，度为 2 节点个数，度为 1 节点个数，以及只有右子节点的节点个数。
5. 已知 $f(x)$ 的原函数为 $F(x)$
 - (a) 如果 $f(x)$ 为奇函数，那么 $F(x)$ 为偶函数
 - (b) 如果 $f(x)$ 为偶函数，那么 $F(x)$ 为奇函数
 - (c) 如果 $f(x)$ 为周期函数，那么 $F(x)$ 为周期函数
 - (d) 如果 $f(x)$ 为增函数，那么 $F(x)$ 为增函数
6. 看代码，判断复杂度：求二叉树高度，选择排序。
7. A 和 B 为任意两事件，下列选项正确的是
 - A. $P(AB) \leq P(A)P(B)$
 - B. $P(AB) \geq P(A)P(B)$
 - C. $P(AB) \leq \frac{P(A)+P(B)}{2}$
 - D. $P(AB) \geq \frac{P(A)+P(B)}{2}$
8. 如何评价一个算法？设计一个火车票系统，需要考虑什么因素？
9. L 为带头节点的链表， p 不为首元素，请排序以下代码删除 p 的直接前驱：
 - (a) $p \rightarrow next = p \rightarrow next \rightarrow next;$
 - (b) $While(p \rightarrow next \rightarrow next != q) p = p \rightarrow next;$
 - (c) $Q = p;$
 - (d) $Free(Q)$
 - (e) $Q = p \rightarrow next;$
 - (f) $p = L;$
10. 已知 A 的特征值为 $-1, 0, 1$ ，对应特征向量为 x_1, x_2, x_3 具体数忘了，求 A^n
11. 判断极限推导过程的正误。

$$\lim_{n \rightarrow +\infty} x^n = L$$

$$x \lim_{n \rightarrow +\infty} x^n = xL$$

$$L = \lim_{n+1 \rightarrow +\infty} x^{n+1} = xL$$

$$(x - 1)L = 0$$

$$\therefore L = 0$$

12. 去除链表 L 头节点，选项不用全选
 - (a) $p \rightarrow next = p \rightarrow next \rightarrow next;$
 - (b) $While(p \rightarrow next \rightarrow next != q) p = p \rightarrow next;$
 - (c) $Q = p;$
 - (d) $Free(Q);$
 - (e) $Q = p \rightarrow next;$
 - (f) $p = L;$
13. 求二叉树叶节点数的代码，给出一个输入样例求出结果。