计算机保研面试_数学常见题

- 计算机保研面试_数学常见题
  - 1 线性代数
    - (1) 余子式
    - (2) 行列式
    - (3) 矩阵的秩，满秩，不满秩
    - (4) 矩阵的迹
    - (5) 线性方程组解的情况
    - (6) 线性相关，线性无关
    - (7) 方阵可逆的充要条件
    - (8) 向量空间，线性空间
    - (9) 向量空间的基和维数
    - (10) 矩阵的特征值和特征向量
    - (11) 向量正交，矩阵正交
    - (12) 正交矩阵
    - (13) 相似矩阵
    - (14) 合同矩阵
    - (15) 正定矩阵
    - (16) 范数
    - (17) 内积
    - (18) 特征值分解
    - (19) 奇异值分解
  - 2 概率论
    - (1) 全概率公式和贝叶斯公式
    - (2) 大数定律
    - (3) 中心极限定理
    - (4) 变量与随机变量有什么区别？
    - (5) 联合概率、边缘概率、条件概率
    - (6) 先验概率、后验概率
    - (7) 常见的概率分布
    - (8) 若干个高斯分布相加/相乘后得到的分布是什么
    - (9) 期望和方差
    - (10) 协方差和相关系数
    - (11) 独立和不相关的区别
    - (12) 独立和互斥的关系
    - (13) 古典概型和几何概型的区别
    - (14) 概率密度函数
    - (15) 极大似然估计
    - (16) 什么是假设和检验
    - (17) 机器学习为什么要使用概率？
    - (18) 不均匀硬币产生等概率
    - (19) 伪随机数产生均匀分布
    - (20) 由均匀分布产生高斯分布
  - 3 微积分
  - 4 离散数学
    - (1) 自反、对称、传递
    - (2) 对称传递能不能推出自反性？
    - (3) 自反性存在有什么意义？
    - (4) 等价关系和等价类
    - (5) 偏序关系

1 线性代数

(1) 余子式

- 余子式：n 阶行列式中，划去 a_{ij} 所在的第 i 行与第 j 列的元，剩下的元所构成的 n-1 阶行列式称为 a_{ij} 的余子式。
- 代数余子式：在余子式前面乘上 $(-1)^{i+j}$

(2) 行列式

- 行列式 det(A)，是一个将方阵 A 映射到实数的函数。
- 几何意义：行列式是有向面积或体积的概念在欧几里得空间中的推广，它描述了一个线性变换对“体积”的影响。
  - 2 阶行列式代表的是平面内的面积。
  - 3 阶行列式代表的是立体空间内的体积。
  - 4 阶行列式是 4 维空间里的超体积。
- 计算方法：行列式等于它任意一行（列）的各元素与其对应的代数余子式乘积之和。
- 行列式和矩阵特征值的关系：行列式等于矩阵特征值的乘积。

(3) 矩阵的秩，满秩，不满秩

- 矩阵的秩
  - 子式角度：矩阵的非零子式的最高阶数。

    > k 阶子式：在矩阵中取 k 行 k 列，交叉处的 k2 个元素按原顺序构成的行列式。

  - 线性无关角度：矩阵的所有行向量中极大线性无关组的元素个数。
  - 标准型角度：先将矩阵转化为行阶梯型，其非零行个数就是矩阵的秩。
- 满秩：一个秩为 n 的矩阵满秩意味着存在一个 n 阶子式不为 0。
- 不满秩：假设其秩为 r，意味着存在一个 r 阶子式不为 0 ，其所有阶数大于 r 的子式都为 0。

(4) 矩阵的迹

方阵 A 的迹：对角线元素之和。

(5) 线性方程组解的情况

- 齐次线性方程组 $AX=0$：$r(A)=n$，有惟一零解；$r(A)<n$，有无穷多解。
- 非齐次线性方程组 $AX=b$：$r(A)\neq r([A,b])$，无解；$r(A)=r([A,b])=n$，有唯一解；$r(A)=r([A,b])<n$，有无穷多解。

(6) 线性相关，线性无关

- 对于线性空间中的 n 个向量，若存在 n 个不全为 0 的常数，使得这 n 个常数与 n 个向量对应乘积加和等于 0，则称这 n 个向量线性相关，如果不存在这样的 n 个常数，称之为线性无关。
- 几何意义：n 个向量线性无关等价于他们所张成的 n 维体的体积不为 0；若线性相关，则体积为 0。

(7) 方阵可逆的充要条件

1. 方阵 A 可逆。
2. A 是非奇异矩阵（$|A|\neq 0$）。
3. A 满秩。
4. A 可以表示成有限个初等矩阵的乘积。
5. 特征值没有 0。
6. 矩阵线性无关。

(8) 向量空间，线性空间

- 向量空间：所有 n 维向量构成的集合称为 n 维向量空间。
- 线性空间：若非空集合 V 上满足加法封闭性和乘法封闭性，并且还满足以下 8 条性质：加法交换律、加法结合律、加法零元、加法逆元、数乘幺元、数乘结合律、数乘分配律（2个），则称集合 V 是线性空间。
- 两者的关系：线性空间包含向量空间，除了向量空间，线性空间还包括多项式空间、矩阵空间等。

(9) 向量空间的基和维数

- 基：在向量空间 V 中可以找到 n 个向量，这 n 个向量线性无关，并且线性空间 V 中的任意一个向量都和这 n 个向量线性相关，那么这 n 个向量就称作线性空间 V 的一个基。
- 维数：基中所含向量个数。

(10) 矩阵的特征值和特征向量

- 含义：矩阵就是一个线性变换，它作用于特征向量后，向量方向保持不变（不发生旋转变换），进行某一比例的伸缩变换，这个比例就是特征值，即

$$AX=\lambda X,\text{其中}X\text{是}A\text{的特征向量},\lambda\text{是}A\text{的特征值}$$

- 特征值和特征向量的关系
  - 特征值和特征向量是一对多的关系，一个特征值可能对应多个特征向量，一个特征向量只能属于一个特征值。
  - 属于不同特征值的特征向量一定线性无关。
  - 设 λ 是 n 阶方阵 A 的一个 k 重特征值（λ 为特征方程的 k 重根），对应于 λ 的线性无关的特征向量的最大个数为 l，则 k≥l，即特征值 λ 的**代数重数不小于几何重数。**
- 应用：奇异值分解、主成分分析。

(11) 向量正交，矩阵正交

- 向量正交：在正交向量组中，任意两个向量的数量积为 0。
- 矩阵正交：两个矩阵正交，表示这两个矩阵相乘结果为单位矩阵。

(12) 正交矩阵

A 是正交矩阵 $\Longleftrightarrow A^TA=E \Longleftrightarrow A^T=A^{-1} \Longleftrightarrow$ A 的各列向量是标准正交向量组（都是单位向量，且两两正交）。

(13) 相似矩阵

- 相似矩阵：对于两个矩阵 A, B，如果存在一个可逆矩阵 P，使得矩阵 $P^{-1}AP=B$，那么矩阵 A 相似于矩阵 B。
- 意义：相似的矩阵是同一个线性变换在不同基/坐标系下的的不同描述。
- 性质：两个相似的矩阵具有相同的秩、迹、行列式、特征值，但特征向量未必相同。
- 补充：实对称矩阵有相同的特征值则必相似。

  > 因为实对称矩阵可对角化，对角化后对角上就是特征值。

(14) 合同矩阵

设 A, B 是两个方阵，若存在可逆矩阵 C，使得 $C^TAC=B$，则称方阵 A, B 合同，记作 $A\simeq B$。

(15) 正定矩阵

- 正定矩阵：假设一个实对称矩阵 A，对于任意一个非零的列向量 S，都有 $S^TAS>0$，则该矩阵正定。
- 半正定矩阵：就是把上面的 $>$ 改成 $\geqslant$
- 性质：
  - 特征值全为正。
  - 各阶主子式大于零。

- 对称矩阵 A 正定的充要条件：A 的特征值全为正。

(16) 范数

- R^n 上的**向量** $x=(x_1,\cdots,x_n)^T$ 的范式（离散型范数公式）：

$$\begin{aligned}&||x||_\infty=\max_{1\le i\le n}|x_i|\\&||x||_1=\sum_{i=1}^n|x_i|\\&||x||_2=(\sum_{i=1}^n x_i^2)^{\frac{1}{2}}\end{aligned}$$

> 带权重的形式（以2-范数为例）： $||x||_2=(\sum_{i=1}^n w_i x_i^2)^{\frac{1}{2}}$

- $C[a,b]$ 上的**函数** $f(x)$ 的范式（连续型范数公式）：

$$\begin{aligned}&||f||_\infty=\max_{a\le x\le b}|f(x)|\\&||f||_1=\int_a^b|f(x)|\mathrm{d}x\\&||f||_2=(\int_a^b f^2(x)\mathrm{d}x)^{\frac{1}{2}}\end{aligned}$$

> 带权重的形式（以2-范数为例）：$||f(x)||_2=(\int_a^b\rho(x)f^2(x)\mathrm{d}x)^{\frac{1}{2}}$

- $R^{n\times n}$ 上的**矩阵** A 的范式:

$$\begin{aligned}&\text{行范式}:\|A\|_\infty=\max_{1\le i\le n}\sum_{j=1}^n|a_{ij}|\\&\text{列范式}:\|A\|_1=\max_{1\le j\le n}\sum_{i=1}^n|a_{ij}|\\&\text{谱范式}:\|A\|_2=\sqrt{\lambda_{max}(A^TA)},\text{其中}\lambda_{max}\text{表示最大特征值}\\&F\text{范式}:\|A\|_F=\sqrt{\sum_{i=1}^n\sum_{j=1}^n a_{ij}^2}\end{aligned}$$

> $A=\begin{bmatrix}4&-3\\-1&6\end{bmatrix}$ 的范式:
>
> $$\begin{aligned}&\text{行范式}:\|A\|_\infty=\max\{\text{各行绝对值之和}\}=\max\{7,7\}=7\\&\text{列范式}:\|A\|_1=\max\{\text{各列绝对值之和}\}=\max\{5,9\}=9\\&F\text{范式}:\|A\|_F=\sqrt{4^2+3^2+1^2+6^2}=7.87401\\&\text{谱范式}:\|A\|_2=\sqrt{\lambda_{max}(\begin{bmatrix}17&-18\\-18&45\end{bmatrix})}=7.335087491\end{aligned}$$
>
> > 怎么求矩阵 A 的最大特征值？ $|\lambda E-A|=0$，然后求 λ 的最大解

(17) 内积

- R^n 上的**向量** x,y 的内积（离散型内积公式）：

$$(x,y)=\sum_{i=1}^n w_i x_i y_i$$

> w_i 是权重

- $C[a,b]$ 上的**函数** $f(x),g(x)$ 的内积（连续型内积公式）：

$$(f(x), g(x)) = \int_a^b \rho(x) f(x) g(x) dx$$

> $\rho(x)$ 是权函数，一般为 1，$\frac{1}{\sqrt{1-x^2}}$，$\sqrt{1-x^2}$

(18) 特征值分解

- 特征值分解（谱分解）：对于一个方阵 A，$\lambda_1 \leqslant \ldots \leqslant \lambda_n$ 是 A 的 n 个特征值，w_1, w_2, w_n 是其对应的 n 个线性无关的特征向量，则存在正交矩阵 W，使得

$$W^{-1}AW = W^TAW = \Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ldots & \\ & & \lambda_n \end{pmatrix}$$

  其中 W 是将 A 的 n 个特征向量单位化后组成的 $n \times n$ 矩阵。

- 意义：将一个方阵对应的线性变换转化为"旋转——拉伸——旋转"变换。

  > 正交矩阵对应旋转变换，对角矩阵对应拉伸变换。

(19) 奇异值分解

奇异值分解（SVD）：特征值分解的扩展，适用于任意矩阵 $A_{m\times n}$。

$$A = U\Sigma V^T$$

- $U_{m\times m}$：对称矩阵，由 AA^T 的所有特征向量组成，其中每个特征向量叫做 A 的左奇异向量。
- $V_{n\times n}$：对称矩阵，由 A^TA 的所有特征向量组成，其中每个特征向量叫做 A 的右奇异向量。
- $\Sigma_{m\times n}$：除了对角线上是奇异值，其他全为 0，奇异值 σ_i 由 $Av_i = \sigma_i u_i$ 求得。

2 概率论

(1) 全概率公式和贝叶斯公式

- 全概率公式
  - 由因推果。造成某种结果，有多种原因，求发生这种结果的概率是多少？
  - 假设造成结果 A 的原因有 $B_1, B_2, \ldots, B_n$，则 A 发生的概率为每个原因发生的概率乘以在这种情况下 A 发生的条件概率的加和，即

$$p(A) = \sum_{i=1}^{n} p(B_\mathrm{i}) \mathrm{p}(\mathrm{A}|\mathrm{B_i})$$

- 贝叶斯公式
  - 由果推因。已知某结果的发生概率，求造成该结果的第 i 个原因的概率是多少？
  - 假设结果 A 发生的概率为 p(A)，原因可能有 $B_1, B_2, \ldots, B_n$，则 A 是由 B_i 造成的概率 $p(B_i|A)$ 这样求

> 先用条件概率公式计算，再用全概率公式替换分母

$$\mathrm{P}\left(\mathrm{B_i}|\mathrm{A}\right)=\frac{\mathrm{P}\left(\mathrm{A}|\mathrm{B_i}\right)\mathrm{P}\left(\mathrm{B}_i\right)}{\mathrm{P}\left(\mathrm{A}\right)}=\frac{\mathrm{P}\left(\mathrm{A}|\mathrm{B_i}\right)\mathrm{P}\left(\mathrm{B}_i\right)}{\sum_{j=1}^{n}P\left(B_j\right)\mathrm{P}\left(\mathrm{A}|\mathrm{B}_j\right)}$$

(2) 大数定律

- 大数定律：当重复试验的次数很大时，随机变量的**均值**依概率收敛于自己的**期望**，“偶然中包含着某种必然”。
- 意义：大数定律将数理统计中的均值和概率论中的期望联系在了一起。
- 三种表述
  - 切比雪夫大数定律：随机变量 $X_1,X_2,...,X_n$ 独立、期望 EX_i 和方差 DX_i 都存在、方差 DX_i 有一致上界（即每个方差都有上界且收敛速度接近），则**样本均值收敛于自己的期望**，即

$$\frac{1}{n}\sum_{i=1}^{n}X_i\xrightarrow{P}\frac{1}{n}\sum_{i=1}^{n}EX_i$$

  - 伯努利大数定律：对于 $X\sim B\left(n,p\right)$，当 n 很大时，有 $\frac{X}{n}\xrightarrow{P}p$。即大量独立重复实验后，随机事件发生的**频率收敛于概率。**
  - 辛钦大数定律：随机变量 $X_1,X_2,...,X_n$ 独立同分布、期望 $EX=\mu$ 存在，当 n 很大时，他们的**算术平均值依概率收敛于期望**，即

$$\frac{1}{n}\sum_{i=1}^{n}X_i\xrightarrow{P}\mu$$

(3) 中心极限定理

- 独立同分布中心极限定理：随机变量 $X_1,X_2,...,X_n$ 独立同分布，期望 $EX=\mu$，方差 $DX=\sigma^2>0$，当 n 很大时，则均值 $\overline{X}_n$ 近似服从正态分布 $N(\mu,\sigma^2/n)$，即

$$\lim_{n\to\infty}P\left(\frac{\overline{X}_n-\mu}{\sigma/\sqrt{n}}<a\right)=\Phi(a)=\int_{-\infty}^{a}\frac{1}{\sqrt{2\pi}}e^{-t^2/2}dt$$

- 拉普拉斯 / 二项分布中心极限定理：是独立同分布中心极限定理的特殊情况，设随机变量 ξ_n 服从二项分布 $B(n,p)$，ξ_n 可视为 n 个独立同分布的 0-1 分布随机变量的和，当 n 很大时，则均值 ξ_n 近似服从正态分布 $N(np,np(1-p))$，即

$$\lim_{n\to\infty}P\left(\frac{\xi_n-np}{\sqrt{np(1-p)}}<a\right)=\Phi(a)$$

> 该定理表明：当试验次数 n 足够大时，二项分布近似于正态分布。

- 应用：如某炮兵阵地对敌人的防御地段进行100次射击，每次射击中炮弹的命中数是一个随机变量，其期望为2，方差为1.69，求在100次射击中有180颗到220颗炮弹命中目标的概率。

(4) 变量与随机变量有什么区别?

- 变量：描述确定性现象，取值固定唯一。
- 随机变量：描述随机现象，取值有多个，且每个取值都有一定的概率。

(5) 联合概率、边缘概率、条件概率

- 联合概率：两个事件共同发生的概率，P(A, B)
- 边缘概率：某个事件发生的概率，与其它事件无关，P(A)
- 条件概率：在事件 B 已经发生的条件下，事件 A 发生的概率，P(A|B)

(6) 先验概率、后验概率

- 先验概率：根据以往经验得到的概率，是"由因推果"问题中的"因"，不需要使用贝叶斯公式计算。
- 后验概率：得到结果后重新修正的概率，是"由果推因"问题中的"因"，需使用贝叶斯公式计算。
- 计算后验概率需要用到先验概率。

(7) 常见的概率分布

分布	分布列 p_k 或分布密度 $p(x)$	期望	方差
$0-1$ 分布 $b(1,p)$	$p_k = p^k(1-p)^{1-k},\quad k=0,1$	p	$p(1-p)$
二项分布 $b(n,p)$ n 重伯努里试验成功事件的次数 有放回抽 n 件，其中一种的件数	$p_k = \binom{n}{k} p^k(1-p)^{n-k},\quad k=0,1,\cdots,n$	np	$np(1-p)$
泊松分布 $P(\lambda)$ 单位时间的发生次数 单位面积 / 体积的数量	$p_k = \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda},\quad k=0,1,\cdots$	λ	λ
超几何分布 $h(n,N,M)$ 不放回抽 n 件，其中一种的件数	$p_k = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}},\quad k=0,1,\cdots,r,\ r=\min\{M,n\}$	$n\frac{M}{N}$	$\frac{nM(N-M)(N-n)}{N^2(N-1)}$
几何分布 $Ge(p)$ 事件首次出现时的试验次数	$p_k = (1-p)^{k-1}p,\quad k=1,2,\cdots$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
负二项分布 $Nb(r,p)$ 事件第 r 次出现时的试验次数	$p_k = \binom{k-1}{r-1}(1-p)^{k-r}p^r,\quad k=r,r+1,\cdots$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$
均匀分布 $U(a,b)$	$p(x) = \frac{1}{b-a},\quad a<x<b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
正态分布 $N(\mu,\sigma^2)$	$p(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}, -\infty<x<\infty$	μ	σ^2
指数分布 $Exp(\lambda)$ 一般为寿命	$p(x) = \lambda\mathrm{e}^{-\lambda x},\quad x\geq 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$

补充几点：

- 0-1 分布又叫伯努利分布、两点分布。
- 泊松分布
  - λ：单位时间 / 面积内随机事件的平均发生次数
  - 泊松分布和二项分布的关系：当二项分布的 n 很大而 p 很小时，泊松分布可作为二项分布的近似
  - 应用：描述单位时间内随机事件发生的次数，如：一天内电路受电磁波干扰的次数
- 超几何分布（不放回抽样）：N 件产品中有 M 件不合格，从 N 件产品中随机抽 n 件检查，发现 k 件不合格品的概率

(8) 若干个高斯分布相加/相乘后得到的分布是什么

假设多个随机变量分别服从不同的高斯分布，如果这些随机变量**彼此独立**，那么这些随机变量的和也服从高斯分布，乘积为高斯分布乘以常数。

> https://blog.csdn.net/qq_41035283/article/details/121015712

(9) 期望和方差

- 期望：随机变量的每个取值与其概率的乘积的累加和，它描述了随机变量的集中特性。
- 方差：随机变量的每个取值减去期望的平方与其概率的乘积的累加和，它描述了随机变量的离散特性。
  - $D(X)=E((X-E(X))^2)=E(X^2)-E^2(X)$ （平方的期望-期望的平方）

(10) 协方差和相关系数

- 协方差：X, Y 的协方差等于每个 X 减去其平均值乘上每个 Y 减去其平均值的乘积的和的平均。

$$\begin{aligned} cov\left(X,Y\right) &= E\left[\left(X-E\left(X\right)\right)\left(Y-E\left(Y\right)\right)\right] \\ &= E\left[XY\right]-E\left[X\right]E\left[Y\right] \end{aligned}$$

  - 意义：协方差表示的是两个变量总体误差的期望，当协方差为 0 时，两者不线性相关；协方差绝对值越大，两者对彼此的影响越大（>0 就是正相关，<0 就是负相关）
  - $var(X+Y)=var(X)+var(Y)+2cov(X+Y)$ ，var 指方差
- 协方差矩阵：n 维随机变量 $X=[X_1,X_2,\ldots,X_n]^T$ 的协方差矩阵为（$c_{ij}=cov(X_i,X_j)$）

$$C=(c_{ij})_{n\times n}=\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}$$

- 相关系数：用 X, Y 的协方差除以 X 的标准差和 Y 的标准差

$$\rho_{X,Y}=\frac{cov\left(X,Y\right)}{\sqrt{D\left(X\right)}\sqrt{D\left(Y\right)}}$$

- 意义：标准化的协方差，消除了两个变量量纲的影响，取值范围为 [0, 1]。

(11) 独立和不相关的区别

- 独立：$P(A|B)=P(A) \Longleftrightarrow P(AB)=P(A)P(B)$
- 不相关：$cov\left(X,Y\right)=E\left[XY\right]-E\left[X\right]E\left[Y\right]=0$
- 独立一定不相关，而不相关不一定独立。**不相关**就是两者**没有线性关系**，但是不排除其它关系存在；**独立**就是**互不相干没有关联**。

(12) 独立和互斥的关系

- 独立：$P(AB)=P(A)P(B)$
- 互斥：$A\cap B=\oslash$
- 如果 $P(A)\neq 0, P(B)\neq 0$，则互斥不独立，独立不互斥。

(13) 古典概型和几何概型的区别

- 古典概型——有限等可能（有限个可能事件，且每个事件都是等可能概率事件）
- 几何概型——无限等可能

(14) 概率密度函数

- 连续型随机变量的取值有无穷多个实数，无法用分布列表示，所以用概率密度函数来表示。
- 概率密度函数不是概率，乘以区间长度微元后就表示概率的近似值。
- 概率密度函数在一段区间上的积分就是随机变量 X 在这段区间上取值的概率。

(15) 极大似然估计

- 一种**参数估计**的方法，它先确定模型，通过若干次实验，观察其结果，反推最有可能（最大概率）导致这样结果的参数值。
- 似然方程的解只是一个估计值，只有在样本数趋于无限多的时候，它才会接近于真实值。
- 步骤：
  1. 写出似然函数
  2. 对似然函数取对数，并整理
  3. 求导数
  4. 解似然方程

(16) 什么是假设和检验

- 假设检验：在总体分布函数未知的情况下，先对总体参数提出一个假设值，然后利用样本信息来判断这一假设是否成立。
- 基本思想：小概率事件在一次的实验中是不可能发生的。
- 两个假设：原假设（要检验的假设）、备选假设（拒绝原假设时的假设），他们是互补的。
- 两类错误
  - 第一类错误 (弃真错误)：原假设 H_0 为真，但是检验出的拒绝原假设，其概率为 α 。
  - 第二类错误 (存伪错误)：原假设 H_0 为假，但是检验出的接受原假设，其概率为 β 。
  - 两类错误是相互矛盾的，不能同时降低两种错误，因此一般在限制 α 的情况下，尽可能地降低 β 的值。
- 基本步骤
  1. 提出原假设和备选假设。
  2. 选取检验统计量。
  3. 给定显著性水平。
  4. 计算检验统计量的具体值，做出判断。

(17) 机器学习为什么要使用概率?

- 机器学习的是由**数据驱动**的方法，它的学习对象是数据，从数据出发提取数据特征抽象出数据模型又从数据中发现知识，最后回到数据的分析和预测中去。
- 机器学习的**算法设计**通常依赖于对数据的概率假设。
- 机器学习模型的训练和预测过程的**评价指标**——模型误差，其本身就是概率的形式。

(18) 不均匀硬币产生等概率

- 已知一随机发生器，产生 0 的概率是 p，产生 1 的概率是 1-p，现在要你构造一个发生器，使得它构造 0 和 1 的概率均为 1/2。
  - 连续随机生成两次，可能的情况有

    00：pp
    01：p(1-p)
    10：(1-p)p
    11：(1-p)(1-p)
  - 所以 2 次 1 组，认为 01 表示 0，10 表示 1，等概率，其他情况放弃。
- 扩展到 3, 4, ..., n 同理，以 3 为例
  - 构造一个发生器，使得它构造 1, 2, 3 的概率均为 1/3：认为 001 表示 1，010 表示 2，100表示 3，等概率，其他情况放弃。

(19) 伪随机数产生均匀分布

混合同余法，其中 m 是模长，a 是乘数，c 是增量，X_0 是初始值。

$$X_n = (aX_{n-1} + b)\%m/m$$

(20) 由均匀分布产生高斯分布

Box-Muller 算法，一共有两种形式，其中 U_1, U_2 是值域为 $(0, 1]$ 的均匀分布的随机数，Z_1, Z_2 是高斯分布的随机数。

$$Z_0 = R\cos(\Theta) = \sqrt{-2\ln U_1}\cos(2\pi U_2)$$

$$Z_1 = R\sin(\Theta) = \sqrt{-2\ln U_1}\sin(2\pi U_2).$$

3 微积分

[保研面试/考研复试高等数学问题整理 - 知乎 (zhihu.com)]

格林公式（Green formula）

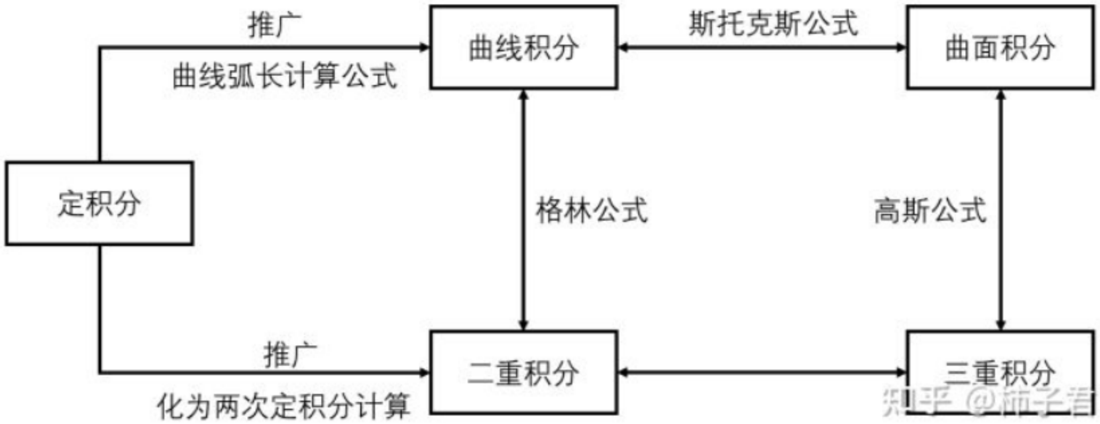
> 若函数 $P(x, y), Q(x, y)$ 在闭区域 D 上连续，且有一阶的连续偏导数，则有 $\iint\limits_D (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})dxdy = \oint_L Pdx + Qdy$，这里 L 为闭区域的边界曲线并取正向。

高斯公式（Gauss formula）

> 设空间闭区域 V 由分片光滑的双侧封闭曲面 S 围城，若函数P, Q, R 在 V 上连续，且有一阶的连续偏导数，则 $\iiint\limits_V (\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z})dxdydz = \oiint\limits_S Pdydz + Qdxdz + Rdxdy$，其中 S 取外侧。

斯托克斯公式（Stokes formula）

> 设光滑曲面 S 的边界 L 是按段光滑的连续曲线，若函数 P, Q, R在 S （连同 L)上连续，且有一阶的连续偏导数，则
> $\oint_L Pdx + Qdy + Rdz = \iint\limits_S (\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z})dydz + (\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x})dzdx$
> $+ (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})dxdy$
> ，其中 S 的侧与 L 的方向按右手法则确定。



4 离散数学

(1) 自反、对称、传递

- 假设集合A，以及基于A上的关系R
  - 自反：　如果a是A的元素，那么<a,a>是R的元素
  - 对称：　如果<a,b>是R的元素，那么<b,a>是R的元素
  - 传递：　如果<a,b>，<b,c>是R的元素，那么<a,c>是R的元素

(2) 对称传递能不能推出自反性?

不能

(3) 自反性存在有什么意义?

自反性是等价关系的基石。

(4) 等价关系和等价类

自反+对称+传递

(5) 偏序关系

自反+反对称+传递