数学基础笔记（V1.01）

你不是一个人在战斗！

haiguang2000@qq.com

最后修改：2018-04-19

目录

机器学习的数学基础

高等数学

1.导数定义：

导数和微分的概念

$$f'(x_0)=\frac{\lim_{\Delta x\to 0} f(x_0+\Delta x)-f(x_0)}{\Delta x} \qquad (1)$$

或者：$$f'(x_0)=\frac{\lim_{x\to x_0} f(x)-f(x_0)}{x-x_0} \qquad (2)$$

2.左右导数导数的几何意义和物理意义

函数$f(x)$在x_0处的左、右导数分别定义为：

左导数：$$f'_-(x_0)=\lim_{\Delta x\to 0^-}\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}=\lim_{x\to x_0^-}\frac{f(x)-f(x_0)}{x-x_0},(x=x_0+\Delta x)$$

右导数：$$f'_+(x_0)=\lim_{\Delta x\to 0^+}\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}=\lim_{x\to x_0^+}\frac{f(x)-f(x_0)}{x-x_0}$$

3.函数的可导性与连续性之间的关系

Th1: 函数$f(x)$在x_0处可微$\Leftrightarrow f(x)$在x_0处可导。

Th2:若函数在点x_0处可导，则$y=f(x)$在点x_0处连续，反之则不成立.即函数连续不一定可导。

Th3:$f'(x_0)$存在$\Leftrightarrow f'_{-¿(x_0)=f'_{+¿(x_0)¿}¿}$

4.平面曲线的切线和法线

切线方程：$y-y_0=f'(x_0)(x-x_0)$

法线方程：$y-y_0=\frac{-1}{f'(x_0)}(x-x_0), f'(x_0)\neq 0$

5.四则运算法则

设函数$u=u(x)$，$v=v(x)$在点x可导，则：

(1) $(u\pm v)'=u'\pm v'$

(2) $(uv)'=uv'+vu'$ $\quad d(uv)=udv+vdu$

(3) $(\frac{u}{v})'=\frac{vu'-uv'}{v^2}(v\neq 0)$ $\quad d(\frac{u}{v})=\frac{vdu-udv}{v^2}$

6.基本导数与微分表

(1) $y=c$（常数） 则：$y'=0$ $\quad dy=0$

(2) $y=x^{\alpha}$(α为实数) 则：$y'=\alpha x^{\alpha-1}$ $\quad dy=\alpha x^{\alpha-1}dx$

(3) $y=a^x$ 则：$y'=a^x\ln a$ $\quad dy=a^x\ln a dx$ 特例: $(e^x)'=e^x$ $\quad d(e^x)=e^x dx$

(4) $y=\log_a x$ 则：

$y'=\frac{1}{x\ln a}$，$dy=\frac{1}{x\ln a}dx$ 特例:$y=\ln x$ $\quad (\ln x)'=\frac{1}{x}$ $\quad d(\ln x)=\frac{1}{x}dx$

(5) $y=\sin x$ 则：$y'=\cos x$ $\quad d(\sin x)=\cos xdx$

(6) $y=\cos x$ 则：$y'=-\sin x$ $\quad d(\cos x)=-\sin xdx$

(7) $y=\tan x$ 则：$y'=\frac{1}{\cos^2 x}=sec^2 x \quad d(\tan x)=sec^2 xdx$

(8) $y=\cot x$ 则：$y'=\frac{-1}{\sin^2 x}=-csc^2 x \quad d(\cot x)=-csc^2 xdx$

(9) $y=sec\,x$ 则：$y'=sec\,x\tan x \quad d(sec\,x)=sec\,x\tan xdx$

(10) $y=csc\,x$ 则：$y'=-csc\,x\cot x \quad d(csc\,x)=-csc\,x\cot xdx$

(11) $y=\arcsin x$ 则：$y'=\frac{1}{\sqrt{1-x^2}} \quad d(\arcsin x)=\frac{1}{\sqrt{1-x^2}}dx$

(12) $y=\arccos x$ 则：$y'=\frac{-1}{\sqrt{1-x^2}} \quad d(\arccos x)=\frac{-1}{\sqrt{1-x^2}}dx$

(13) $y=\arctan x$ 则：$y'=\frac{1}{1+x^2} \quad d(\arctan x)=\frac{1}{1+x^2}dx$

(14) $y=\text{arccot}\, x$ 则：$y'=\frac{-1}{1+x^2} \quad d(\text{arccot}\, x)=\frac{-1}{1+x^2}dx$

(15) $y=shx$ 则：$y'=chx \quad d(shx)=chxdx$

(16) $y=chx$ 则：$y'=shx \quad d(chx)=shxdx$

7.复合函数，反函数，隐函数以及参数方程所确定的函数的微分法

(1) 反函数的运算法则: 设$y=f(x)$在点x的某邻域内单调连续，在点x处可导且$f'(x)\neq 0$，

则其反函数在点x所对应的y处可导，并且有$\frac{dy}{dx}=\frac{1}{\frac{dx}{dy}}$

(2) 复合函数的运算法则:若$\mu=\varphi(x)$在点x可导,而$y=f(\mu)$在对应点μ($\mu=\varphi(x)$)可导,则复合函数$y=f(\varphi(x))$在点x可导,且$y'=f'(\mu)\cdot\varphi'(x)$

(3) 隐函数导数$\frac{dy}{dx}$的求法一般有三种方法：

1)方程两边对x求导，要记住y是x的函数，则y的函数是x的复合函数.例如$\frac{1}{y}$，y^2，lny，e^y等均是x的复合函数. 对x求导应按复合函数连锁法则做。

2)公式法.由$F(x,y)=0$知$\frac{dy}{dx}=\frac{-F'_x(x,y)}{F'_y(x,y)}$,其中，$F'_x(x,y)$，$F'_y(x,y)$分别表示$F(x,y)$对$x$和$y$的偏导数。

3)利用微分形式不变性

8.常用高阶导数公式

（1）$(a^x)^{(n)}=a^x\ln^n a(a>0)(e^x)^{(n)}=e^x$

（2）$(\sin kx)^{(n)}=k^n\sin(kx+n\cdot\frac{\pi}{2})$

（3）$(\cos kx)^{(n)}=k^n\cos(kx+n\cdot\frac{\pi}{2})$

（4）$(x^m)^{(n)}=m(m-1)\cdots(m-n+1)x^{m-n}$

（5）$(\ln x)^{(n)}=(-1)^{(n-1)}\frac{(n-1)!}{x^n}$

（6）莱布尼兹公式：若$u(x),v(x)$均n阶可导，则：$(uv)^{(n)}=\sum_{i=0}^{n}c_n^i u^{(i)}v^{(n-i)}$，其中$u^{(0)}=u$，$v^{(0)}=v$

9.微分中值定理，泰勒公式

Th1:(费马定理)

若函数$f(x)$满足条件：

(1)函数$f(x)$在x_0的某邻域内有定义，并且在此邻域内恒有$f(x)\le f(x_0)$或$f(x)\ge f(x_0)$,

(2) $f(x)$在x_0处可导,则有 $f'(x_0)=0$

Th2:(罗尔定理)

设函数$f(x)$满足条件：

(1)在闭区间$[a,b]$上连续； (2)在(a,b)内可导；(3)$f(a)=f(b)$

则在(a,b)内∃一个ξ，使 $f'(\xi)=0$

Th3: (拉格朗日中值定理)

设函数$f(x)$满足条件：

(1)在$[a,b]$上连续；(2)在(a,b)内可导；

则在(a,b)内存在一个ξ，使 $\frac{f(b)-f(a)}{b-a}=f'(\xi)$

Th4: (柯西中值定理)

设函数$f(x)$，$g(x)$满足条件：

(1) 在$[a,b]$上连续；(2) 在(a,b)内可导且$f'(x)$，$g'(x)$均存在，且$g'(x)\neq 0$

则在(a,b)内存在一个ξ，使 $\frac{f(b)-f(a)}{g(b)-g(a)}=\frac{f'(\xi)}{g'(\xi)}$

10.洛必达法则

法则Ⅰ($\frac{0}{0}$型不定式极限)

设函数$f(x),g(x)$满足条件： $\lim\limits_{x\to x_0} f(x)=0, \lim\limits_{x\to x_0} g(x)=0$; $f(x),g(x)$在x_0的邻域内可导 (在x_0处可除外)且$g'(x)\neq 0$;

$\frac{\lim\limits_{x\to x_0} f'(x)}{g'(x)}$存在(或∞)。

则：$\dfrac{\lim\limits_{x\to x_0} f(x)}{g(x)}=\dfrac{\lim\limits_{x\to x_0} f'(x)}{g'(x)}$

法则I’($\dfrac{0}{0}$型不定式极限)

设函数$f(x),g(x)$满足条件：$\lim\limits_{x\to\infty} f(x)=0,\lim\limits_{x\to\infty} g(x)=0$;存在一个$X>0$,当$|x|>X$时,

$f(x),g(x)$可导,且$g'(x)\neq 0$;$\dfrac{\lim\limits_{x\to x_0} f'(x)}{g'(x)}$存在(或∞)。

则：$\dfrac{\lim\limits_{x\to x_0} f(x)}{g(x)}=\dfrac{\lim\limits_{x\to x_0} f'(x)}{g'(x)}$.

法则Ⅱ($\dfrac{\infty}{\infty}$型不定式极限)

设函数$f(x),g(x)$满足条件：$\lim\limits_{x\to x_0} f(x)=\infty,\lim\limits_{x\to x_0} g(x)=\infty$; $f(x),g(x)$在x_0 的邻域内可 导

(在x_0处可除外)且$g'(x)\neq 0$;$\dfrac{\lim\limits_{x\to x_0} f'(x)}{g'(x)}$存在(或∞)。

则：$\dfrac{\lim\limits_{x\to x_0} f(x)}{g(x)}=\dfrac{\lim\limits_{x\to x_0} f'(x)}{g'(x)}$.

同理法则II’($\dfrac{\infty}{\infty}$型不定式极限)仿法则I’可写出

11.泰勒公式

设函数$f(x)$在点x_0处的某邻域内具有$n+1$阶导数，则对该邻域内异于x_0的任意点x，在x_0与x之间至少存在一个ξ，使得：

$$f(x)=f(x_0)+f'(x_0)(x-x_0)+\frac{1}{2!}f''(x_0)(x-x_0)^2+\cdots\frac{+f^{(n)}(x_0)}{n!}(x-x_0)^n+R_n(x)$$

其中 $R_n(x)=\frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}$称为$f(x)$在点$x_0$处的$n$阶泰勒余项。

令$x_0=0$，则n阶泰勒公式：

$$f(x)=f(0)+f'(0)x+\frac{1}{2!}f''(0)x^2+\cdots+\frac{f^{(n)}(0)}{n!}x^n+R_n(x)\ldots\ldots$$

(1) 其中 $R_n(x)=\frac{f^{(n+1)}(\xi)}{(n+1)!}x^{n+1}$，$\xi$在 0 与$x$之间。(1)式称为麦克劳林公式

常用五种函数在$x_0=0$处的泰勒公式 ：

1) $e^x=1+x+\frac{1}{2!}x^2+\cdots+\frac{1}{n!}x^n+\frac{x^{n+1}}{(n+1)!}e^{\xi}$

或¿$1+x+\frac{1}{2!}x^2+\cdots+\frac{1}{n!}x^n+o(x^n)$

2) $\sin x=x-\frac{1}{3!}x^3+\cdots+\frac{x^n}{n!}\sin\frac{n\pi}{2}+\frac{x^{n+1}}{(n+1)!}\sin\left(\xi+\frac{n+1}{2}\pi\right)$

或¿$x-\frac{1}{3!}x^3+\cdots+\frac{x^n}{n!}\sin\frac{n\pi}{2}+o(x^n)$

3) $\cos x=1-\frac{1}{2!}x^2+\cdots+\frac{x^n}{n!}\cos\frac{n\pi}{2}+\frac{x^{n+1}}{(n+1)!}\cos(\xi+\frac{n+1}{2}\pi)$

或¿$1-\frac{1}{2!}x^2+\cdots+\frac{x^n}{n!}\cos\frac{n\pi}{2}+o(x^n)$

4) $\ln(1+x)=x-\frac{1}{2}x^2+\frac{1}{3}x^3-\cdots+(-1)^{n-1}\frac{x^n}{n}+\frac{(-1)^n x^{n+1}}{(n+1)(1+\xi)^{n+1}}$

或¿$x-\frac{1}{2}x^2+\frac{1}{3}x^3-\cdots+(-1)^{n-1}\frac{x^n}{n}+o(x^n)$

5) $(1+x)^m=1+mx+\frac{m(m-1)}{2!}x^2+\cdots+\frac{m(m-1)\cdots(m-n+1)}{n!}x^n$

$+\frac{m(m-1)\cdots(m-n+1)}{(n+1)!}x^{n+1}(1+\xi)^{m-n-1}$

或 $(1+x)^m=1+mx+\frac{m(m-1)}{2!}x^2+\cdots+\frac{m(m-1)\cdots(m-n+1)}{n!}x^n+o(x^n)$

12.函数单调性的判断

Th1: 设函数$f(x)$在(a,b)区间内可导，如果对$\forall x\in(a,b)$，都有$f'(x)>0$（或$f'(x)<0$），则函数$f(x)$在(a,b)内是单调增加的（或单调减少）。

Th2:（取极值的必要条件）设函数$f(x)$在x_0处可导，且在x_0处取极值，则$f'(x_0)=0$.

Th3:（取极值的第一充分条件）设函数$f(x)$在x_0的某一邻域内可微，且$f'(x_0)=0$（或$f(x)$在x_0处连续，但$f'(x_0)$不存在.）。

(1)若当x经过x_0时，$f'(x)$由“+”变“-”，则$f(x_0)$为极大值；

(2)若当x经过x_0时，$f'(x)$由“-”变“+”，则$f(x_0)$为极小值；

(3)若$f'(x)$经过$x=x_0$的两侧不变号，则$f(x_0)$不是极值。

Th4: (取极值的第二充分条件)设$f(x)$在点x_0处有$f''(x)\neq 0$，且$f'(x_0)=0$，则：

当$f''(x_0)<0$时，$f(x_0)$为极大值；当$f''(x_0)>0$时，$f(x_0)$为极小值. 注：如果$f''(x_0)0$，此方法失效。

13.渐近线的求法

(1)水平渐近线

若$\lim\limits_{x\to+\infty}f(x)=b$，或$\lim\limits_{x\to-\infty}f(x)=b$，则$y=b$ 称为函数$y=f(x)$的水平渐近线。

(2)铅直渐近线

若$\lim\limits_{x \to x_0^{-¿} f(x)=\infty ¿}$ ¿，或$\lim\limits_{x \to x_0^{+¿} f(x)=\infty ¿}$ ¿，则$x=x_0$称为$y=f(x)$的铅直渐近线。

(3)斜渐近线 若$a=\dfrac{\lim\limits_{x \to \infty} f(x)}{x}, b=\lim\limits_{x \to \infty}[f(x)-ax]$，则 $y=ax+b$称为$y=f(x)$的斜渐近线。

14.函数凹凸性的判断

Th1: (凹凸性的判别定理）若在 I 上$f''(x)<0$（或$f''(x)>0$），则$f(x)$在 I 上是凸的（或凹的）。

Th2: (拐点的判别定理 1)若在x_0处$f''(x)=0$，（或$f''(x)$不存在），当x变动经过x_0时，$f''(x)$变号，则$(x_0, f(x_0))$为拐点。

Th3: (拐点的判别定理 2)设$f(x)$在x_0点的某邻域内有三阶导数，且$f''(x)=0$，$f'''(x)\neq 0$，则$(x_0, f(x_0))$为拐点。

15.弧微分

$$dS=\sqrt{1+y'^2}dx$$

16.曲率

曲线$y=f(x)$在点(x,y)处的曲率$k=\dfrac{|y''|}{(1+y'^2)^{3/2}}$. 对于参数方程：

$$\begin{cases} x=\varphi(t) \\ y=\psi(t) \end{cases}, k=\frac{|\varphi'(t)\psi''(t)-\varphi''(t)\psi'(t)|}{[\varphi'^2(t)+\psi'^2(t)]^{3/2}}$$

17.曲率半径

曲线在点M处的曲率$k(k\neq 0)$与曲线在点M处的曲率半径ρ有如下关系：$\rho=\dfrac{1}{k}$

线性代数

行列式

1.行列式按行（列）展开定理

(1) 设$A=\left(a_{ij}\right)_{n\times n}$，则：$a_{i1}A_{j1}+a_{i2}A_{j2}+\cdots+a_{\text{¿}}A_{jn}=\begin{cases}|A|, i=j\\ 0, i\neq j\end{cases}$

或$a_{1i}A_{1j}+a_{2i}A_{2j}+\cdots+a_{\text{¿}}A_{nj}=\begin{cases}|A|, i=j\\ 0, i\neq j\end{cases}$

即 $AA^{\text{¿}}=A^{\text{¿}}A=|A|E$,其中：$A^{\text{¿}}=\begin{pmatrix}A_{11} & A_{12} & \ldots & A_{1n}\\ A_{21} & A_{22} & \ldots & A_{2n}\\ \ldots & \ldots & \ldots & \ldots\\ A_{n1} & A_{n2} & \ldots & A_{nn}\end{pmatrix}=\left(A_{ji}\right)=\left(A_{ij}\right)^{T}$

$$D_n=|\begin{matrix}1 & 1 & \ldots & 1\\ x_1 & x_2 & \ldots & x_n\\ \ldots & \ldots & \ldots & \ldots\\ x_1^{n-1} & x_2^{n-1} & \ldots & x_n^{n-1}\end{matrix}|=\text{¿}\prod_{1\leq j<i\leq n}\square\left(x_i-x_j\right)$$

(2) 设A,B为n阶方阵，则$|AB|=|A||B|=|B||A|=|BA|$，但$|A\pm B|=|A|\pm|B|$不一定成立。

(3) $|kA|=k^n|A|,A$为n阶方阵。

(4) 设A为n阶方阵，$\text{¿}A^T\vee\text{¿}\vee A\vee;\vee A^{-1}\vee\text{¿}\vee A\text{¿}^{-1}$（若$A$可逆），$\text{¿}A^{\text{¿}}\vee\text{¿}\vee A\text{¿}^{n-1}$

$n\geq 2$

(5) $\begin{vmatrix}A & O\\ O & B\end{vmatrix}=\begin{vmatrix}A & C\\ O & B\end{vmatrix}=\begin{vmatrix}A & O\\ C & B\end{vmatrix}=\text{¿}$ ，A,B为方阵，但$\begin{vmatrix}O & A_{m\times m}\\ B_{n\times n} & O\end{vmatrix}=\text{¿}$。

(6) 范德蒙行列式$D_n=\begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ \cdots & \cdots & \cdots & \cdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{vmatrix}=\prod_{1\le j<i\le n}(x_i-x_j)$

设A是n阶方阵，$\lambda_i(i=1,2\cdots;n)$是A的n个特征值，则$|A|=\prod_{i=1}^{n}\lambda_i$

矩阵

矩阵：$m\times n$个数a_{ij}排成m行n列的表格$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$称为矩阵，简记为$A$，或者$(a_{ij})_{m\times n}$。

若$m=n$，则称A是n阶矩阵或n阶方阵。

矩阵的线性运算

1.矩阵的加法

设$A=(a_{ij}),B=(b_{ij})$是两个$m\times n$矩阵，则$m\times n$ 矩阵$C=(c_{ij})=a_{ij}+b_{ij}$称为矩阵A与B的和，记为$A+B=C$。

2.矩阵的数乘

设$A=(a_{ij})$是$m\times n$矩阵，k是一个常数，则$m\times n$矩阵$(k a_{ij})$称为数k与矩阵A的数乘，记为kA。

3.矩阵的乘法

设$A=(a_{ij})$是$m\times n$矩阵，$B=(b_{ij})$是$n\times s$矩阵，那么$m\times s$矩阵$C=(c_{ij})$，其中

$c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{¿}b_{nj}=\sum_{k=1}^{n}a_{ik}b_{kj}$称为$AB$的乘积，记为$C=AB$。

4. A^T、A^{-1}、$A^{¿}$三者之间的关系

(1) $(A^T)^T=A,(AB)^T=B^T A^T,(kA)^T=k A^T,(A\pm B)^T=A^T\pm B^T$

(2) $(A^{-1})^{-1}=A,(AB)^{-1}=B^{-1}A^{-1},(kA)^{-1}=\frac{1}{k}A^{-1},$

但$(A\pm B)^{-1}=A^{-1}\pm B^{-1}$不一定成立。

(3) $(A^{¿})^{¿}=¿A¿^{n-2}A(n\geq 3)$，$(AB)^{¿}=B^{¿}A^{¿}$，$(kA)^{¿}=k^{n-1}A^{¿}(n\geq 2)$

但$(A\pm B)^{¿}=A^{¿}\pm B^{¿}$不一定成立。

(4) $(A^{-1})^T=(A^T)^{-1},(A^{-1})^{¿}=(AA^{¿})^{-1},(A^{¿})^T=(A^T)^{¿}$

5.有关$A^{¿}$的结论

(1) $AA^{¿}=A^{¿}A=¿A\vee E$

(2) $¿A^{¿}\vee¿\vee A¿^{n-1}(n\geq 2),(kA)^{¿}=k^{n-1}A^{¿},(A^{¿})^{¿}=¿A¿^{n-2}A(n\geq 3)$

(3) 若A可逆，则$A^{¿}=¿A\vee A^{-1},(A^{¿})^{¿}=\frac{1}{¿A\vee¿A¿}$

(4) 若A为n阶方阵，则：

$$r(A^{¿})=\begin{cases} n,r(A)=n \\ 1,r(A)=n-1 \\ 0,r(A)<n-1 \end{cases}$$

6.有关A^{-1}的结论

A可逆$\Leftrightarrow AB=E;\Leftrightarrow \vee A\vee \neq 0;\Leftrightarrow r(A)=n;$

$\Leftrightarrow A$可以表示为初等矩阵的乘积；$\Leftrightarrow A$无零特征值 ；$\Leftrightarrow Ax=0$只有零解。

7.有关矩阵秩的结论

(1) 秩$r(A)$=行秩=列秩；

(2) $r(A_{m\times n})\leq min(m,n);$

(3) $A\neq 0\Rightarrow r(A)\geq 1$；

(4) $r(A\pm B)\leq r(A)+r(B);$

(5) 初等变换不改变矩阵的秩

(6) $r(A)+r(B)-n\leq r(AB)\leq min(r(A),r(B))$,特别若$AB=O$

则：$r(A)+r(B)\leq n$

(7) 若A^{-1}存在$\Rightarrow r(AB)=r(B);$ 若B^{-1}存在 $\Rightarrow r(AB)=r(A);$

若$r(A_{m\times n})=n\Rightarrow r(AB)=r(B);$ 若$r(A_{m\times s})=n\Rightarrow r(AB)=r(A)$。

(8) $r(A_{m\times s})=n\Leftrightarrow Ax=0$只有零解

8.分块求逆公式

$$\begin{pmatrix} A & O \\ O & B \end{pmatrix}^{-1}=\begin{pmatrix} A^{-1} & O \\ O & B^{-1} \end{pmatrix};\qquad \begin{pmatrix} A & C \\ O & B \end{pmatrix}^{-1}=\begin{pmatrix} A^{-1}-A^{-1}CB^{-1} \\ OB^{-1} \end{pmatrix};$$

$$\begin{pmatrix} A & O \\ C & B \end{pmatrix}^{-1}=\begin{pmatrix} A^{-1}O \\ -B^{-1}CA^{-1}B^{-1} \end{pmatrix};\qquad \begin{pmatrix} O & A \\ B & O \end{pmatrix}^{-1}=\begin{pmatrix} O & B^{-1} \\ A^{-1} & O \end{pmatrix}$$

这里A，B均为可逆方阵。

向量

1.有关向量组的线性表示

(1) $\alpha_1,\alpha_2,\cdots;\alpha_s$线性相关⇔至少有一个向量可以用其余向量线性表示。

(2) $\alpha_1,\alpha_2,\cdots;\alpha_s$线性无关，$\alpha_1,\alpha_2,\cdots;\alpha_s$，$\beta$线性相关⇔$\beta$可以由$\alpha_1,\alpha_2,\cdots;\alpha_s$唯一线性表示。

(3) β可以由$\alpha_1,\alpha_2,\cdots;\alpha_s$线性表示 ⇔$r(\alpha_1,\alpha_2,\cdots;\alpha_s)=r(\alpha_1,\alpha_2,\cdots;\alpha_s,\beta)$。

2.有关向量组的线性相关性

(1)部分相关，整体相关；整体无关，部分无关.

(2) ① n个n维向量 $\alpha_1,\alpha_2\cdots\alpha_n$线性无关⇔$\left|[\alpha_1\alpha_2\cdots\alpha_n]\right|\neq 0$， n个n维向量$\alpha_1,\alpha_2\cdots\alpha_n$线性相关 ⇔$\vee[\alpha_1,\alpha_2,\cdots;\alpha_n]\vee¿0$。

② $n+1$个n维向量线性相关。

③ 若$\alpha_1,\alpha_2\cdots\alpha_S$线性无关，则添加分量后仍线性无关；或一组向量线性相关，去掉某些分量后仍线性相关。

3.有关向量组的线性表示

(1) $\alpha_1,\alpha_2,\cdots;\alpha_s$线性相关⇔至少有一个向量可以用其余向量线性表示。

(2) $\alpha_1,\alpha_2,\cdots;\alpha_s$线性无关，$\alpha_1,\alpha_2,\cdots;\alpha_s$，$\beta$线性相关⇔$\beta$ 可以由$\alpha_1,\alpha_2,\cdots;\alpha_s$唯一线性表示。

(3) β可以由$\alpha_1,\alpha_2,\cdots;\alpha_s$线性表示 ⇔$r(\alpha_1,\alpha_2,\cdots;\alpha_s)=r(\alpha_1,\alpha_2,\cdots;\alpha_s,\beta)$

4.向量组的秩与矩阵的秩之间的关系

设$r(A_{m\times n})=r$，则A的秩$r(A)$与A的行列向量组的线性相关性关系为：

(1) 若$r\left(A_{m\times n}\right)=r=m$，则$A$的行向量组线性无关。

(2) 若$r\left(A_{m\times n}\right)=r<m$，则$A$的行向量组线性相关。

(3) 若$r\left(A_{m\times n}\right)=r=n$，则$A$的列向量组线性无关。

(4) 若$r\left(A_{m\times n}\right)=r<n$，则$A$的列向量组线性相关。

5.n维向量空间的基变换公式及过渡矩阵

若$\alpha_1,\alpha_2,\cdots;\alpha_n$与$\beta_1,\beta_2,\cdots;\beta_n$是向量空间$V$的两组基，则基变换公式为：

$$\left(\beta_1,\beta_2,\cdots;\beta_n\right)=\left(\alpha_1,\alpha_2,\cdots;\alpha_n\right)\begin{bmatrix} c_{11}\,c_{12}\cdots c_{1n} \\ c_{21}\,c_{22}\cdots c_{2n} \\ \cdots\cdots\cdots\cdots \\ c_{n1}\,c_{n2}\cdots c_{nn} \end{bmatrix}=\left(\alpha_1,\alpha_2,\cdots;\alpha_n\right)C$$

其中C是可逆矩阵，称为由基$\alpha_1,\alpha_2,\cdots;\alpha_n$到基$\beta_1,\beta_2,\cdots;\beta_n$的过渡矩阵。

6.坐标变换公式

若向量γ在基$\alpha_1,\alpha_2,\cdots;\alpha_n$与基$\beta_1,\beta_2,\cdots;\beta_n$的坐标分别是 $X=\left(x_1,x_2,\cdots;x_n\right)^T$，

$Y=\left(y_1,y_2,\cdots;y_n\right)^T$ 即：$\gamma=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n=y_1\beta_1+y_2\beta_2+\cdots+y_n\beta_n$，则向量坐标变换公式为$X=CY$ 或 $Y=C^{-1}X$ ，其中C是从基$\alpha_1,\alpha_2,\cdots;\alpha_n$到基$\beta_1,\beta_2,\cdots;\beta_n$的过渡矩阵。

7.向量的内积

$\left(\alpha,\beta\right)=a_1b_1+a_2b_2+\cdots+a_nb_n=\alpha^T\beta=\beta^T\alpha$

8.Schmidt 正交化

若$\alpha_1,\alpha_2,\cdots;\alpha_s$线性无关，则可构造$\beta_1,\beta_2,\cdots;\beta_s$使其两两正交，且$\beta_i$仅是$\alpha_1,\alpha_2,\cdots;\alpha_i$的线性组合$(i=1,2,\cdots;n)$，再把$\beta_i$单位化，记$\gamma_i=\frac{\beta_i}{|\beta_i|}$，则$\gamma_1,\gamma_2,\cdots;\gamma_i$是规范正交向量组。其中

$$\beta_1=\alpha_1, \quad \beta_2=\alpha_2-\frac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1, \quad \beta_3=\alpha_3-\frac{(\alpha_3,\beta_1)}{(\beta_1,\beta_1)}\beta_1-\frac{(\alpha_3,\beta_2)}{(\beta_2,\beta_2)}\beta_2,$$

…………

$$\beta_s=\alpha_s-\frac{(\alpha_s,\beta_1)}{(\beta_1,\beta_1)}\beta_1-\frac{(\alpha_s,\beta_2)}{(\beta_2,\beta_2)}\beta_2-\cdots-\frac{(\alpha_s,\beta_{s-1})}{(\beta_{s-1},\beta_{s-1})}\beta_{s-1}$$

9.正交基及规范正交基

向量空间一组基中的向量如果两两正交，就称为正交基；若正交基中每个向量都是单位向量，就称其为规范正交基。

线性方程组

1．克莱姆法则

线性方程组$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=b_n \end{cases}$，如果系数行列式$D=|A|\neq 0$，则方程组有唯一解，$x_1=\frac{D_1}{D},x_2=\frac{D_2}{D},\cdots;x_n=\frac{D_n}{D}$，其中$D_j$是把$D$中第$j$列元素换成方程组右端的常数列所得的行列式。

2. n阶矩阵A可逆$\Leftrightarrow Ax=0$只有零解。$\Leftrightarrow \forall b, Ax=b$总有唯一解，一般地，$r(A_{m\times n})=n\Leftrightarrow Ax=0$只有零解。

3.非奇次线性方程组有解的充分必要条件，线性方程组解的性质和解的结构

(1) 设A为$m\times n$矩阵，若$r(A_{m\times n})=m$，则对$Ax=b$而言必有$r(A)=r(A\ \vdots\ b)=m$，从而$Ax=b$有解。

(2) 设$x_1,x_2,\cdots x_s$为$Ax=b$的解，则$k_1x_1+k_2x_2+\cdots+k_sx_s$当$k_1+k_2+\cdots+k_s=1$时仍为$Ax=b$的解；但当$k_1+k_2+\cdots+k_s=0$时，则为$Ax=0$的解。特别$\frac{x_1+x_2}{2}$为$Ax=b$的解；$2x_3-(x_1+x_2)$为$Ax=0$的解。

(3) 非齐次线性方程组$Ax=b$无解$\Leftrightarrow r(A)+1=r(\overline{A})\Leftrightarrow b$不能由$A$的列向量$\alpha_1,\alpha_2,\cdots,\alpha_n$线性表示。

4.奇次线性方程组的基础解系和通解，解空间，非奇次线性方程组的通解

(1) 齐次方程组$Ax=0$恒有解(必有零解)。当有非零解时，由于解向量的任意线性组合仍是该齐次方程组的解向量，因此$Ax=0$的全体解向量构成一个向量空间，称为该方程组的解空间，解空间的维数是$n-r(A)$，解空间的一组基称为齐次方程组的基础解系。

(2) $\eta_1,\eta_2,\cdots,\eta_t$是$Ax=0$的基础解系，即：

1) $\eta_1,\eta_2,\cdots,\eta_t$是$Ax=0$的解；

2) $\eta_1,\eta_2,\cdots,\eta_t$线性无关；

3) $Ax=0$的任一解都可以由$\eta_1,\eta_2,\cdots,\eta_t$线性表出. $k_1\eta_1+k_2\eta_2+\cdots+k_t\eta_t$是$Ax=0$的通解，其中$k_1,k_2,\cdots,k_t$是任意常数。

矩阵的特征值和特征向量

1.矩阵的特征值和特征向量的概念及性质

(1) 设λ是A的一个特征值，则 $kA, aA+bE, A^2, A^m, f(A), A^T, A^{-1}, A^{¿}$有一个特征值分别为 $k\lambda, a\lambda+b, \lambda^2, \lambda^m, f(\lambda), \lambda, \lambda^{-1}, ¿ A \vee \frac{¿}{\lambda}, ¿$且对应特征向量相同（$A^T$ 例外）。

(2) 若$\lambda_1, \lambda_2, \cdots; \lambda_n$为$A$的$n$个特征值，则$\sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} a_{ii}, \prod_{i=1}^{n} \lambda_i = ¿ A \vee ¿$,从而$¿ A \vee \neq 0 \Leftrightarrow A$没有特征值。

(3) 设$\lambda_1, \lambda_2, \cdots; \lambda_s$为$A$的$s$个特征值，对应特征向量为 $\alpha_1, \alpha_2, \cdots; \alpha_s$，

若: $\alpha = k_1 \alpha_1 + k_2 \alpha_2 + \cdots + k_s \alpha_s$，

则: $A^n \alpha = k_1 A^n \alpha_1 + k_2 A^n \alpha_2 + \cdots + k_s A^n \alpha_s = k_1 \lambda_1^n \alpha_1 + k_2 \lambda_2^n \alpha_2 + \cdots k_s \lambda_s^n \alpha_s$。

2.相似变换、相似矩阵的概念及性质

(1) 若$A \sim B$，则

1) $A^T \sim B^T, A^{-1} \sim B^{-1}, , A^{¿} \sim B^{¿}$

2) $¿ A \vee ¿ \vee B \vee, \sum_{i=1}^{n} A_{ii} = \sum_{i=1}^{n} b_{ii}, r(A) = r(B)$

3) $¿ \lambda E - A \vee ¿ \vee \lambda E - B \vee ¿$，对$\forall$ λ成立

3.矩阵可相似对角化的充分必要条件

(1) 设A为n阶方阵，则A可对角化$\Leftrightarrow$对每个k_i重根特征值λ_i，有$n - r(\lambda_i E - A) = k_i$

(2) 设A可对角化，则由$P^{-1} AP = \Lambda$,有$A = P \Lambda P^{-1}$，从而$A^n = P \Lambda^n P^{-1}$

(3) 重要结论

1) 若$A \sim B, C \sim D$，则$\begin{bmatrix} A & O \\ O & C \end{bmatrix} \sim \begin{bmatrix} B & O \\ O & D \end{bmatrix}$.

2) 若$A\sim B$，则$f(A)\sim f(B),|f(A)|\sim|f(B)|$，其中$f(A)$为关于$n$阶方阵$A$的多项式。

3) 若A为可对角化矩阵，则其非零特征值的个数(重根重复计算) = 秩(A)

4.实对称矩阵的特征值、特征向量及相似对角阵

(1)相似矩阵：设A,B为两个n阶方阵，如果存在一个可逆矩阵P，使得$B=P^{-1}AP$成立，则称矩阵A与B相似，记为$A\sim B$。

(2)相似矩阵的性质：如果$A\sim B$则有：

1) $A^T\sim B^T$

2) $A^{-1}\sim B^{-1}$（若A，B均可逆）

3) $A^k\sim B^k$（k为正整数）

4) $|\lambda E-A|=|\lambda E-B|$，从而$A,B$ 有相同的特征值

5) $|A|=|B|$，从而A,B同时可逆或者不可逆

6) 秩$(A)=¿$秩$(B),|\lambda E-A|=|\lambda E-B|$，$A,B$不一定相似

二次型

1.n个变量$x_1,x_2,\cdots;x_n$的二次齐次函数

$f(x_1,x_2,\cdots;x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_iy_j$，其中$a_{ij}=a_{ji}(i,j=1,2,\cdots;n)$，称为$n$元二次型，简称二

次型. 若令$x=\begin{bmatrix}x_1\\x_1\\\vdots\\x_n\end{bmatrix},A=\begin{bmatrix}a_{11}a_{12}\cdots a_{1n}\\a_{21}a_{22}\cdots a_{2n}\\\cdots\cdots\cdots\\a_{n1}a_{n2}\cdots a_{nn}\end{bmatrix}$,这二次型$f$可改写成矩阵向量形式$f=x^TAx$。其中$A$

称为二次型矩阵，因为$a_{ij}=a_{ji}(i,j=1,2,\cdots;n)$，所以二次型矩阵均为对称矩阵，且二次型与对称矩阵一一对应，并把矩阵A的秩称为二次型的秩。

2.惯性定理，二次型的标准形和规范形

(1) 惯性定理

对于任一二次型，不论选取怎样的合同变换使它化为仅含平方项的标准型，其正负惯性指数与所选变换无关，这就是所谓的惯性定理。

(2) 标准形

二次型$f=(x_1,x_2,\cdots;x_n)=x^T Ax$经过合同变换$x=Cy$化为$f=x^T Ax=y^T C^T AC$

$y=\sum_{i=1}^{r} d_i y_i^2$称为 $f(r\leq n)$的标准形。在一般的数域内，二次型的标准形不是唯一的，与所作的合同变换有关，但系数不为零的平方项的个数由$r(A$的秩$)$唯一确定。

(3) 规范形

任一实二次型f都可经过合同变换化为规范形$f=z_1^2+z_2^2+\cdots+z_p^2-z_{p+1}^2-\cdots-z_r^2$，其中$r$为$A$的秩，$p$为正惯性指数，$r-p$为负惯性指数，且规范型唯一。

3.用正交变换和配方法化二次型为标准形，二次型及其矩阵的正定性

设A正定$\Rightarrow kA(k>0),A^T,A^{-1},A^{¿}$正定；$¿A\vee¿0,A$可逆；$a_{ii}>0$，且$¿A_{ii}\vee¿0$

A，B正定$\Rightarrow A+B$正定，但AB，BA不一定正定

A正定$\Leftrightarrow f(x)=x^T Ax>0,\forall x\neq 0$

$\Leftrightarrow A$的各阶顺序主子式全大于零

$\Leftrightarrow A$的所有特征值大于零

$\Leftrightarrow A$的正惯性指数为n

$\Leftrightarrow$存在可逆阵P使$A=P^T P$

$\Leftrightarrow$存在正交矩阵Q，使$Q^T AQ=Q^{-1} AQ=\begin{pmatrix} \lambda_1 & ¿ ¿ \lambda_n ¿ \\ ¿ & ¿ \end{pmatrix}$,

其中$\lambda_i>0, i=1,2,\cdots;n$.正定$\Rightarrow kA(k>0), A^T, A^{-1}, A^{¿}$正定；$¿A \vee ¿0, A$可逆；$a_{ii}>0$，且$¿A_{ii} \vee ¿0$。

概率论和数理统计

随机事件和概率

1.事件的关系与运算

(1) 子事件：$A \subset B$，若A发生，则B发生。

(2) 相等事件：$A=B$，即$A \subset B$，且$B \subset A$ 。

(3) 和事件：$A \bigcup B$（或$A+B$），A与B中至少有一个发生。

(4) 差事件：$A-B$，A发生但B不发生。

(5) 积事件：$A \bigcap B$（或AB），A与B同时发生。

(6) 互斥事件（互不相容）：$A \bigcap B=\varnothing$。

(7) 互逆事件（对立事件）： $A \bigcap B=\varnothing, A \bigcup B=\Omega, A=\overline{B}, B=\overline{A}$ 。

2.运算律

(1) 交换律：$A \bigcup B=B \bigcup A, A \bigcap B=B \bigcap A$

(2) 结合律：$(A \bigcup B) \bigcup C=A \bigcup (B \bigcup C)$；$(A \bigcap B) \bigcap C=A \bigcap (B \bigcap C)$

(3) 分配律：$(A \bigcup B) \bigcap C=(A \bigcap C) \bigcup (B \bigcap C)$

3.德.摩根律

$\overline{A \bigcup B}=\overline{A} \bigcap \overline{B} \quad \overline{A \bigcap B}=\overline{A} \bigcup \overline{B}$

4.完全事件组

$A_1 A_2 \cdots A_n$两两互斥，且和事件为必然事件，即$A_i \bigcap A_j=\varnothing, i \neq j, \bigcup_{i=1}^{n}=\Omega$

5.概率的基本概念

(1) 概率：事件发生的可能性大小的度量，其严格定义如下：

概率$P(g)$为定义在事件集合上的满足下面 3 个条件的函数：

1)对任何事件A，$P(A)\geq 0$

2)对必然事件Ω，$P(\Omega)=1$

3)对$A_1 A_2 \cdots A_n, \cdots$,若$A_i A_j=\varnothing (i\neq j)$，则：$P(\bigcup_{i=1}^{\infty} A_i)=\sum_{i=1}^{\infty} P(A).$

(2) 概率的基本性质

1) $P(\overline{A})=1-P(A);$

2) $P(A-B)=P(A)-P(AB);$

3) $P(A\bigcup B)=P(A)+P(B)-P(AB)$ 特别，当$B\subset A$时，$P(A-B)=P(A)-P(B)$且$P(B)\leq P(A)$；

$P(A\bigcup B\bigcup C)=P(A)+P(B)+P(C)-P(AB)-P(BC)-P(AC)+P(ABC)$ 4) 若

$A_1, A_2, \cdots; A_n$两两互斥，则$P(\bigcup_{i=1}^{n} A_i)=\sum_{i=1}^{n}$ ¿¿

(3) 古典型概率: 实验的所有结果只有有限个， 且每个结果发生的可能性相同，其概率计算公式： $P(A)=\frac{\text{事件}A\text{发生的基本事件数}}{\text{基本事件总数}}$

(4) 几何型概率: 样本空间Ω为欧氏空间中的一个区域， 且每个样本点的出现具有等可能性，其概率计算公式：$P(A)=\frac{A\text{的度量(长度、 面积、 体积)}}{\Omega\text{的度量(长度、 面积、 体积)}}$

6.概率的基本公式

(1) 条件概率: $P(B \vee A)=\frac{P(AB)}{P(A)}$,表示A发生的条件下，B发生的概率

(2) 全概率公式： $P(A)=\sum_{i=1}^{n} P(A \vee B_i)P(B_i), B_i B_j=\varnothing, i\neq j, \bigcup_{i=1}^{n} B_i=\Omega.$

(3) **Bayes** 公式：

$$P(B_j \vee A)=\frac{P(A \vee B_j)P(B_j)}{\sum_{i=1}^{n} P(A \vee B_i)P(B_i)}, j=1,2,\cdots;n$$

注：上述公式中事件B_i的个数可为可列个.

(4)乘法公式： $P(A_1 A_2)=P(A_1)P(A_2 \vee A_1)=P(A_2)P(A_1 \vee A_2)$

$$P(A_1 A_2 \cdots A_n)=P(A_1)P(A_2 \vee A_1)P(A_3 \vee A_1 A_2)\cdots P(A_n \vee A_1 A_2 \cdots A_{n-1})$$

7.事件的独立性

(1) A 与 B 相互独立⇔ $P(AB)=P(A)P(B)$

(2) A，B，C 两两独立 ⇔ $P(AB)=P(A)P(B); P(BC)=P(B)P(C);$

$P(AC)=P(A)P(C);$

(3) A，B，C 相互独立 ⇔ $P(AB)=P(A)P(B); P(BC)=P(B)P(C);$

$P(AC)=P(A)P(C); P(ABC)=P(A)P(B)P(C).$

8.独立重复试验

将某试验独立重复 n 次，若每次实验中事件 A 发生的概率为 p，则 n 次试验中 A 发生 k 次

的概率为： $P(X=k)=C_n^k p^k (1-p)^{n-k}$。

9.重要公式与结论

(1) $P(\overline{A})=1-P(A)$

(2) $P(A \bigcup B)=P(A)+P(B)-P(AB)$

$P(A \bigcup B \bigcup C)=P(A)+P(B)+P(C)-P(AB)-P(BC)-P(AC)+P(ABC)$

(3) $P(A-B)=P(A)-P(AB)$

(4) $P(A\overline{B})=P(A)-P(AB), P(A)=P(AB)+P(A\overline{B})$,

$P(A \bigcup B)=P(A)+P(\overline{A}B)=P(AB)+P(A\overline{B})+P(\overline{A}B)$

(5) 条件概率$P(\cdot \vee B)$满足概率的所有性质，

例如：. $P(\overline{A}_1 \vee B)=1-P(A_1 \vee B)$

$P(A_1 \bigcup A_2 \vee B)=P(A_1 \vee B)+P(A_2 \vee B)-P(A_1 A_2 \vee B)$

$P(A_1 A_2 \vee B)=P(A_1 \vee B)P(A_2 \vee A_1 B)$

(6) 若$A_1, A_2, \cdots; A_n$相互独立，则$P(\text{¿} i=1 \text{¿} n\, A_i)=\prod_{i=1}^{n} P(A_i)$,

$P(\text{¿} i=1 \text{¿} n\, A_i)=\prod_{i=1}^{n} (1-P(A_i))$

(7) 互斥、互逆与独立性之间的关系： A 与 B 互逆⇒A 与 B 互斥，但反之不成立，A 与 B 互 斥（或互逆）且均非零概率事件⇒A 与 B 不独立.

(8) 若$A_1, A_2, \cdots; A_m, B_1, B_2, \cdots; B_n$相互独立，则$f(A_1, A_2, \cdots; A_m)$与 $g(B_1, B_2, \cdots; B_n)$也相互独立，其中$f(\cdot), g(\cdot)$分别表示对相应事件做任意事件运算后所得的事件，另外，概率为 1（或 0）的事件与任何事件相互独立.

随机变量及其概率分布

1.随机变量及概率分布

取值带有随机性的变量，严格地说是定义在样本空间上，取值于实数的函数称为随机变量，概率分布通常指分布函数或分布律

2.分布函数的概念与性质

定义：$F(x)=P(X\leq x), -\infty<x<+\infty$

性质：(1)$0\leq F(x)\leq 1$　　(2)$F(x)$单调不减

(3)右连续$F(x+0)=F(x)$　　(4)$F(-\infty)=0, F(+\infty)=1$

3.离散型随机变量的概率分布

$$P(X=x_i)=p_i, i=1,2,\cdots,n,\cdots p_i\geq 0, \sum_{i=1}^{\infty} p_i=1$$

4.连续型随机变量的概率密度

概率密度$f(x)$;非负可积，且:(1)$f(x)\geq 0$, (2)$\int_{-\infty}^{+\infty} f(x)dx=1$ (3)x为$f(x)$的连续点，则:

$f(x)=F'(x)$分布函数$F(x)=\int_{-\infty}^{x} f(t)dt$

5.常见分布

(1) 0-1 分布:$P(X=k)=p^k(1-p)^{1-k}, k=0,1$

(2) 二项分布:$B(n,p)$： $P(X=k)=C_n^k p^k(1-p)^{n-k}, k=0,1,\cdots,n$

(3) **Poisson** 分布:$p(\lambda)$： $P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}, \lambda>0, k=0,1,2\cdots$

(4) 均匀分布$U(a,b)$：$f(x)=\begin{cases}\frac{1}{b-a}, a<x<b\\ 0,\end{cases}$

(5) 正态分布:$N(\mu,\sigma^2)$: $\varphi(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{\frac{-(x-\mu)^2}{2\sigma^2}},\sigma>0,-\infty<x<+\infty$

(6)指数分布:$E(\lambda)$:$f(x)=\begin{cases}\lambda e^{-\lambda x}, x>0, \lambda>0\\ 0,\end{cases}$

(7)几何分布:$G(p)$:$P(X=k)=(1-p)^{k-1}p, 0<p<1, k=1,2,\cdots$:

(8)超几何分布: $H(N,M,n)$:$P(X=k)=\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}, k=0,1,\cdots;min(n,M)$

6.随机变量函数的概率分布

(1)离散型：$P(X=x_1)=p_i, Y=g(X)$

则: $P(Y=y_j)=\sum_{g(x_i)=y_i}P(X=x_i)$

(2)连续型：$X\sim f_X(x), Y=g(x)$

则:$F_y(y)=P(Y\le y)=P(g(X)\le y)=\int_{g(x)\le y}f_x(x)dx$，$f_Y(y)=F'_Y(y)$

7.重要公式与结论

(1) $X \sim N(0,1)\Rightarrow\varphi(0)=\frac{1}{\sqrt{2\pi}},\Phi(0)=\frac{1}{2},\Phi(-a)=P(X\le -a)=1-\Phi(a)$

(2) $X \sim N(\mu,\sigma^2)\Rightarrow\frac{X-\mu}{\sigma} \sim N(0,1), P(X\le a)=\Phi(\frac{a-\mu}{\sigma})$

(3) $X \quad E(\lambda) \Rightarrow P(X>s+t \vee X>s)=P(X>t)$

(4) $X \quad G(p) \Rightarrow P(X=m+k \vee X>m)=P(X=k)$

(5) 离散型随机变量的分布函数为阶梯间断函数；连续型随机变量的分布函数为连续函数，但不一定为处处可导函数。

(6) 存在既非离散也非连续型随机变量。

多维随机变量及其分布

1.二维随机变量及其联合分布

由两个随机变量构成的随机向量(X,Y)，联合分布为$F(x,y)=P(X\le x,Y\le y)$

2.二维离散型随机变量的分布

(1) 联合概率分布律 $P\{X=x_i,Y=y_j\}=p_{ij};i,j=1,2,\cdots$

(2) 边缘分布律 $p_{i\cdot}=\sum_{j=1}^{\infty} p_{ij},i=1,2,\cdots$ $p_{\cdot j}=\sum_{i}^{\infty} p_{ij},j=1,2,\cdots$

(3) 条件分布律 $P\{X=x_i \vee Y=y_j\}=\frac{p_{ij}}{p_{\cdot j}}$ $\quad P\{Y=y_j \vee X=x_i\}=\frac{p_{ij}}{p_{i\cdot}}$

3. 二维连续性随机变量的密度

(1) 联合概率密度$f(x,y)$:

1) $f(x,y)\ge 0$ 2) $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dxdy=1$

(2) 分布函数：$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)dudv$

(3) 边缘概率密度： $f_X(x)=\int_{-\infty}^{+\infty} f(x,y)dy$ $f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)dx$

(4) 条件概率密度：$f_{X\vee Y}(x|y)=\frac{f(x,y)}{f_Y(y)}$ $f_{Y\vee X}(y\vee x)=\frac{f(x,y)}{f_X(x)}$

4.常见二维随机变量的联合分布

(1) 二维均匀分布：$(x,y)\sim U(D)$,$f(x,y)=\begin{cases}\frac{1}{S(D)},(x,y)\in D\\ 0,\text{其他}\end{cases}$

(2) 二维正态分布：$(X,Y ¿ N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}.\exp\left\{\frac{-1}{2(1-\rho^2)}[\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}]\right\}$$

5.随机变量的独立性和相关性

X和Y的相互独立:$\Leftrightarrow F(x,y)=F_X(x)F_Y(y)$:

$\Leftrightarrow p_{ij}=p_{i\cdot}\cdot p_{\cdot j}$（离散型） $\Leftrightarrow f(x,y)=f_X(x)f_Y(y)$（连续型）

X和Y的相关性：

相关系数$\rho_{XY}=0$时，称X和Y不相关， 否则称X和Y相关

6.两个随机变量简单函数的概率分布

离散型： $P(X=x_i,Y=y_i)=p_{ij},Z=g(X,Y)$ 则：

$$P(Z=z_k)=P\{g(X,Y)=z_k\}=\sum_{g(x_i,y_i)=z_k} P\{X=x_i,Y=y_j\}$$

连续型：$(X,Y)\sim f(x,y),Z=g(X,Y)$ 则：

$$F_z(z)=P\{g(X,Y)\le z\}=\iint\limits_{g(x,y)\le z} f(x,y)dxdy$$，$f_z(z)=F'_z(z)$

7.重要公式与结论

(1) 边缘密度公式：$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)dy,\ f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)dx$

(2) $P\{(X,Y)\in D\}=\iint\limits_{D} f(x,y)dxdy$

(3) 若(X,Y)服从二维正态分布$N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$ 则有：

1) $X \quad N(\mu_1,\sigma_1^2), Y \quad N(\mu_2,\sigma_2^2).$

2) X与Y相互独立$\Leftrightarrow \rho=0$，即X与Y不相关。

3) $C_1X+C_2Y \quad N(C_1\mu_1+C_2\mu_2, C_1^2\sigma_1^2+C_2^2\sigma_2^2+2C_1C_2\sigma_1\sigma_2\rho)$

4) X关于 Y=y 的条件分布为：$N(\mu_1+\rho\frac{\sigma_1}{\sigma_2}(y-\mu_2),\sigma_1^2(1-\rho^2))$

5) Y关于$X=x$的条件分布为：$N(\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1),\sigma_2^2(1-\rho^2))$

(4) 若X与Y独立，且分别服从$N(\mu_1,\sigma_1^2),N(\mu_1,\sigma_2^2)$，则：

$(X,Y) \quad N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,0)$，$C_1X+C_2Y \sim N(C_1\mu_1+C_2\mu_2, C_1^2\sigma_1^2+C_2^2\sigma_2^2).$

(5) 若X与Y相互独立，$f(x)$和$g(x)$为连续函数， 则$f(X)$和$g(Y)$也相互独立。

随机变量的数字特征

1.数学期望

离散型：$P\{X=x_i\}=p_i, E(X)=\sum_i x_i p_i$；

连续型： $X \quad f(x), E(X)=\int_{-\infty}^{+\infty} xf(x)dx$

性质：

(1) $E(C)=C, E[E(X)]=E(X)$

(2) $E(C_1X+C_2Y)=C_1E(X)+C_2E(Y)$

(3) 若X和Y独立，则$E(XY)=E(X)E(Y)$ (4)$[E(XY)]^2\leq E(X^2)E(Y^2)$

2.方差：$D(X)=E[X-E(X)]^2=E(X^2)-[E(X)]^2$

3.标准差：$\sqrt{D(X)}$，

4.离散型：$D(X)=\sum_i [x_i-E(X)]^2 p_i$

5.连续型：$D(X)=\int_{-\infty}^{+\infty}[x-E(X)]^2 f(x)dx$

性质：

(1)$D(C)=0, D[E(X)]=0, D[D(X)]=0$

(2)X与Y相互独立，则$D(X\pm Y)=D(X)+D(Y)$

(3)$D(C_1X+C_2)=C_1^2D(X)$

(4) 一般有 $D(X\pm Y)=D(X)+D(Y)\pm 2Cov(X,Y)=D(X)+D(Y)\pm 2\rho\sqrt{D(X)}\sqrt{D(Y)}$

(5)$D(X)<E(X-C)^2, C\neq E(X)$

(6)$D(X)=0\Leftrightarrow P\{X=C\}=1$

6.随机变量函数的数学期望

(1) 对于函数$Y=g(x)$

X为离散型：$P\{X=x_i\}=p_i, E(Y)=\sum_i g(x_i)p_i$；

X为连续型：$X\quad f(x), E(Y)=\int_{-\infty}^{+\infty} g(x)f(x)dx$

(2) $Z=g(X,Y);(X,Y)\quad P\{X=x_i, Y=y_j\}=p_{ij}$; $E(Z)=\sum_i\sum_j g(x_i,y_j)p_{ij}$

$(X,Y)\quad f(x,y); E(Z)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)f(x,y)dxdy$

7.协方差　$Cov(X,Y)=E$¿

8.相关系数 $\rho_{XY}=\frac{Cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$,$k$阶原点矩 $E(X^k)$; k阶中心矩 $E\{[X-E(X)]^k\}$

性质：

(1)$Cov(X,Y)=Cov(Y,X)$

(2)$Cov(aX,bY)=abCov(Y,X)$

(3)$Cov(X_1+X_2,Y)=Cov(X_1,Y)+Cov(X_2,Y)$

(4) $\left|\rho(X,Y)\right|\le 1$

(5) $\rho(X,Y)=1\Leftrightarrow P(Y=aX+b)=1$ ，其中$a>0$

$\rho(X,Y)=-1\Leftrightarrow P(Y=aX+b)=1$ ，其中$a<0$

9.重要公式与结论

(1) $D(X)=E(X^2)-E^2(X)$

(2) $Cov(X,Y)=E(XY)-E(X)E(Y)$

(3) $\left|\rho(X,Y)\right|\le 1$,且 $\rho(X,Y)=1\Leftrightarrow P(Y=aX+b)=1$，其中$a>0$

$\rho(X,Y)=-1\Leftrightarrow P(Y=aX+b)=1$，其中$a<0$

(4) 下面 5 个条件互为充要条件：

$\rho(X,Y)=0\Leftrightarrow Cov(X,Y)=0\Leftrightarrow E(X,Y)=E(X)E(Y)\Leftrightarrow D(X+Y)=D(X)+D(Y)$
$\Leftrightarrow D(X-Y)=D(X)+D(Y)$

注：X与Y独立为上述 5 个条件中任何一个成立的充分条件，但非必要条件。

数理统计的基本概念

1.基本概念

总体：研究对象的全体，它是一个随机变量，用X表示。

个体：组成总体的每个基本元素。

简单随机样本：来自总体X的n个相互独立且与总体同分布的随机变量$X_1,X_2\cdots;X_n$，称为容量为n的简单随机样本，简称样本。

统计量：设$X_1, X_2 \cdots; X_n$,是来自总体X的一个样本，$g\left(X_1, X_2 \cdots; X_n\right)$）是样本的连续函数，且$g(\bullet)$中不含任何未知参数，则称$g\left(X_1, X_2 \cdots; X_n\right)$为统计量

样本均值：$\overline{X}=\frac{1}{n}\sum_{i=1}^{n} X_i$

样本方差：$S^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$

样本矩：样本k阶原点矩：$A_k=\frac{1}{n}\sum_{i=1}^{n} X_i^k, k=1,2,\cdots$

样本k阶中心矩：$B_k=\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^k, k=1,2,\cdots$

2.分布

χ^2分布：$\chi^2=X_1^2+X_2^2+\cdots+X_n^2 \quad \chi^2(n)$，其中$X_1, X_2 \cdots; X_n$,相互独立，且同服从$N(0,1)$

t分布：$T=\frac{X}{\sqrt{Y/n}} \quad t(n)$ ，其中$X \quad N(0,1), Y \quad \chi^2(n)$,且$X$，$Y$ 相互独立。

F 分布：$F=\frac{X/n_1}{Y/n_2} \quad F\left(n_1, n_2\right)$，其中$X \quad \chi^2\left(n_1\right), Y \quad \chi^2\left(n_2\right)$,且$X$，$Y$相互独立。

分位数：若$P\left(X \leq x_\alpha\right)=\alpha$,则称$x_\alpha$为$X$的$\alpha$分位数

3.正态总体的常用样本分布

(1) 设$X_1, X_2 \cdots; X_n$为来自正态总体$N\left(\mu, \sigma^2\right)$的样本，

$\overline{X}=\frac{1}{n}\sum_{i=1}^{n} X_i, S^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$,则：

1) $\overline{X} \quad N\left(\mu,\frac{\sigma^2}{n}\right)$或者$\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \quad N(0,1)$

2) $\frac{(n-1)S^2}{\sigma^2}=\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\overline{X})^2 \quad \chi^2(n-1)$

3) $\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2 \quad \chi^2(n)$

4)$\frac{\overline{X}-\mu}{S/\sqrt{n}} \quad t(n-1)$

4.重要公式与结论

(1) 对于$\chi^2 \quad \chi^2(n)$，有$E(\chi^2(n))=n,D(\chi^2(n))=2n;$

(2) 对于$T \quad t(n)$，有$E(T)=0,D(T)=\frac{n}{n-2}(n>2)$；

(3) 对于$F \square F(m,n)$，有 $\frac{1}{F} \quad F(n,m),F_{a/2}(m,n)=\frac{1}{F_{1-a/2}(n,m)};$

(4) 对于任意总体X，有 $E(\overline{X})=E(X),E(S^2)=D(X),D(\overline{X})=\frac{D(X)}{n}$