**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

（1）

或者： （2）

**2.左右导数-导数的几何意义和物理意义**

函数在处的左、右导数分别定义为：

左导数：

右导数：

**3.函数的可导性与连续性之间的关系**

**Th1:** 函数在处可微在处可导。

**Th2:**若函数在点处可导，则在点处连续，反之则不成立。即函数连续不一定可导。

**Th3:**存在

**4.平面曲线的切线和法线**

切线方程 :

法线方程：

**5.四则运算法则**

设函数在点可导，则：

(1)

(2)

(3)

**6.基本导数与微分表**

(1) （常数） 则：

(2) (为实数) 则：

(3) 则： 特例:

(4) 则：

特例:

(5) 则：

(6) 则：

(7) 则：

(8) 则：

(9) 则：

(10) 则：

(11) 则：

(12) 则：

(13) 则：

(14) 则：

(15) 则：

(16) 则：

**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: 设在点的某邻域内单调连续，在点处可导且，则其反函数在点所对应的处可导，并且有

(2) 复合函数的运算法则:若在点可导,而在对应点()可导,则复合函数在点可导,且

(3) 隐函数导数的求法一般有三种方法：

1)方程两边对求导，要记住是的函数，则的函数是的复合函数.例如，，，等均是的复合函数. 对求导应按复合函数连锁法则做。

2)公式法.由知 ,其中，， 分别表示对和的偏导数。

3)利用微分形式不变性

**8.常用高阶导数公式**

（1）

（2）

（3）

（4）

（5）

（6）莱布尼兹公式：若均阶可导，则： ，其中，

**9.微分中值定理，泰勒公式**

**Th1:**(费马定理)

若函数满足条件：

(1)函数在的某邻域内有定义，并且在此邻域内恒有 或,

(2) 在处可导,则有

**Th2:**(罗尔定理)

设函数满足条件：

(1)在闭区间上连续； (2)在内可导；(3)

则在内一个，使

**Th3:** (拉格朗日中值定理)

设函数满足条件：

(1)在上连续；(2)在内可导；

则在内存在一个，使

**Th4:** (柯西中值定理)

设函数，满足条件：

(1) 在上连续；(2) 在内可导且，均存在，且

则在内存在一个，使

**10.洛必达法则**

**法则Ⅰ(型不定式极限)**

设函数满足条件： ; 在的邻域内可导 (在处可除外)且;

存在(或)。

则：

**法则 (型不定式极限)**

设函数满足条件： ;存在一个,当时,可导,且;存在(或)。

则：

**法则Ⅱ(型不定式极限)**

设函数满足条件： ; 在 的邻域内可 导(在处可除外)且;存在(或)。

则：

同理法则(型不定式极限)仿法则可写出

**11.泰勒公式**

设函数在点处的某邻域内具有阶导数，则对该邻域内异于的任意点，在与之间至少存在一个，使得：

其中 称为在点处的阶泰勒余项。

令，则阶泰勒公式：

……

1. 其中 ，在0与之间。(1)式称为麦克劳林公式

常用五种函数在处的泰勒公式 ：

1)   
 或

2)

或

3)

或

4)

或

5)

或

**12.函数单调性的判断**

**Th1:** 设函数在区间内可导，如果对，都有（或），则函数在内是单调增加的（或单调减少）。

**Th2:** （取极值的必要条件）设函数在处可导，且在处取极值，则.

**Th3:** （取极值的第一充分条件）设函数在的某一邻域内可微，且（或在处连续，但不存在.）。

(1)若当经过时，由“+”变“-”，则为极大值；

(2)若当经过时，由“-”变“+”，则为极小值；

(3)若经过的两侧不变号，则不是极值。

Th4: (取极值的第二充分条件)设在点处有，且，则：

当时，为极大值； 当时，为极小值. 注：如果，此方法失效。

**13.渐近线的求法**

(1)水平渐近线

若，或，则 称为函数的水平渐近线。

(2)铅直渐近线

若，或，则 称为的铅直渐近线。

(3)斜渐近线 若，则 称为的斜渐近线。

**14.函数凹凸性的判断**

**Th1:** (凹凸性的判别定理）若在I上（或）， 则在I上是凸的（或凹的）。

**Th2:** (拐点的判别定理1)若在处，（或不存在），当变动经过时，变号，则为拐点。

**Th3:** (拐点的判别定理2)设在点的某邻域内有三阶导数，且，，则为拐点。

**15.弧微分**

**16.曲率**

曲线在点处的曲率 对于参数方程：

**17.曲率半径**

曲线在点处的曲率与曲线在点处的曲率半径有如下关系：

**线性代数**

**行列式**

**1.行列式按行（列）展开定理**

(1) 设，则：

或

即 其中：

(2) 设为阶方阵，则，但不一定成立。

(3) ,为阶方阵。

(4) 设为阶方阵，（若可逆），

(5) ，为方阵，但 。

(6) 范德蒙行列式

设是阶方阵，是的个特征值，则

**矩阵**

矩阵：个数排成行列的表格称为矩阵，简记为，或者 。若，则称是阶矩阵或阶方阵。

**矩阵的线性运算**

**1.矩阵的加法**

设是两个矩阵，则 矩阵称为矩阵与的和，记为 。

**2.矩阵的数乘**

设是矩阵，是一个常数，则矩阵称为数与矩阵的数乘，记为。

**3.矩阵的乘法**

设是矩阵，是矩阵，那么矩阵，其中 称为的乘积，记为 。

**4. 、、三者之间的关系**

(1)

(2)

但 不一定成立。

(3) ，

但不一定成立。

(4)

**5.有关的结论**

(1)

(2)

(3) 若可逆，则

(4) 若为阶方阵，则：

**6.有关的结论**

可逆

可以表示为初等矩阵的乘积；。

**7.有关矩阵秩的结论**

(1) 秩=行秩=列秩；

(2)

(3) ；

(4)

(5) 初等变换不改变矩阵的秩

(6) 特别若

则：

(7) 若存在 若存在

若 若。

(8) 只有零解

**8.分块求逆公式**

； ；

；

这里，均为可逆方阵。

**向量**

**1.有关向量组的线性表示**

(1) 线性相关至少有一个向量可以用其余向量线性表示。

(2) 线性无关，，线性相关可以由唯一线性表示。

(3) 可以由线性表示 。

**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 个维向量 线性无关， 个维向量线性相关 。

② 个维向量线性相关。

③ 若线性无关，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) 线性相关至少有一个向量可以用其余向量线性表示。

(2) 线性无关，，线性相关 可以由唯一线性表示。

(3) 可以由线性表示

**4.向量组的秩与矩阵的秩之间的关系**

设，则的秩与的行列向量组的线性相关性关系为：

(1) 若，则的行向量组线性无关。

(2) 若，则的行向量组线性相关。

(3) 若，则的列向量组线性无关。

(4) 若，则的列向量组线性相关。

**5.维向量空间的基变换公式及过渡矩阵**

若与是向量空间的两组基，则基变换公式为：

其中是可逆矩阵，称为由基到基的过渡矩阵。

**6.坐标变换公式**

若向量在基与基的坐标分别是 ，

即： ，则向量坐标变换公式为 或 ，其中是从基到基的过渡矩阵。

**7.向量的内积**

**8.Schmidt正交化**

若线性无关，则可构造使其两两正交，且仅是的线性组合，再把单位化，记，则是规范正交向量组。其中 ， ， ，

............

**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，如果系数行列式，则方程组有唯一解， ，其中是把中第列元素换成方程组右端的常数列所得的行列式。

**2.** 阶矩阵可逆只有零解。总有唯一解，一般地，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) 设为矩阵，若，则对而言必有，从而有解。

(2) 设为的解，则当时仍为的解；但当时，则为的解。特别为的解；为的解。

(3) 非齐次线性方程组无解不能由的列向量线性表示。

**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，解空间的一组基称为齐次方程组的基础解系。

(2) 是的基础解系，即：

1) 是的解；

2) 线性无关；

3) 的任一解都可以由线性表出. 是的通解，其中是任意常数。

**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) 设是的一个特征值，则 有一个特征值分别为 且对应特征向量相同（ 例外）。

(2) 若为的个特征值，则 ,从而没有特征值。

(3) 设为的个特征值，对应特征向量为 ，

若: ,

则: 。

**2.相似变换、相似矩阵的概念及性质**

(1) 若，则

1)

2)

3) ，对成立

**3.矩阵可相似对角化的充分必要条件**

(1) 设为阶方阵，则可对角化对每个重根特征值，有

(2) 设可对角化，则由有，从而

(3) 重要结论

1) 若，则.

2) 若，则，其中为关于阶方阵的多项式。

3) 若为可对角化矩阵，则其非零特征值的个数(重根重复计算)＝秩()

**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：设为两个阶方阵，如果存在一个可逆矩阵，使得成立，则称矩阵与相似，记为。

(2)相似矩阵的性质：如果则有：

1)

2) （若，均可逆）

3) （为正整数）

4) ，从而 有相同的特征值

5) ，从而同时可逆或者不可逆

6) 秩秩，不一定相似

**二次型**

**1.个变量的二次齐次函数**

，其中，称为元二次型，简称二次型. 若令,这二次型可改写成矩阵向量形式。其中称为二次型矩阵，因为，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

二次型经过合同变换化为

称为 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唯一确定。

(3) 规范形

任一实二次型都可经过合同变换化为规范形，其中为的秩，为正惯性指数，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**

设正定正定；,可逆；，且

，正定正定，但，不一定正定

正定

的各阶顺序主子式全大于零

的所有特征值大于零

的正惯性指数为

存在可逆阵使

存在正交矩阵，使

其中正定正定； 可逆；，且 。

**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：，若发生，则发生。

(2) 相等事件：，即，且 。

(3) 和事件：（或），与中至少有一个发生。

(4) 差事件：，发生但不发生。

(5) 积事件：（或），与同时发生。

(6) 互斥事件（互不相容）：=。

(7) 互逆事件（对立事件）： 。

2.运算律

(1) 交换律：

(2) 结合律：；

(3) 分配律：

**3.德摩根律**

**4.完全事件组**

两两互斥，且和事件为必然事件，即

**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，

2)对必然事件，

3)对 ,若，则：

(2) 概率的基本性质

1) ;

2)

3) 特别，当时，且； 4) 若两两互斥，则

(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：

(4) 几何型概率: 样本空间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：

**6.概率的基本公式**

(1) 条件概率: ,表示发生的条件下，发生的概率

(2) 全概率公式：

(3) **Bayes**公式：

注：上述公式中事件的个数可为可列个.

(4)乘法公式：

**7.事件的独立性**

(1) A与B相互独立

(2) A，B，C两两独立

(3) A，B，C相互独立

**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： 。

**9.重要公式与结论**

(1)

(2)

(3)

(4)

(5) 条件概率满足概率的所有性质，

例如：.

(6) 若相互独立，则

(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) 若相互独立，则与 也相互独立，其中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：

性质：(1) (2)单调不减

(3)右连续 (4)

**3.离散型随机变量的概率分布**

**4.连续型随机变量的概率密度**

概率密度非负可积，且:(1) (2) (3)为的连续点，则:

分布函数

**5.常见分布**

(1) 0-1分布:

(2) 二项分布:：

(3) **Poisson**分布:：

(4) 均匀分布：

(5) 正态分布:

(6)指数分布:

(7)几何分布:

(8)超几何分布:

**6.随机变量函数的概率分布**

(1)离散型：

则:

(2)连续型：

则:，

**7.重要公式与结论**

(1)

(2)

(3)

(4)

(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， 联合分布为

**2.二维离散型随机变量的分布**

(1) 联合概率分布律

(2) 边缘分布律

(3) 条件分布律

**3. 二维连续性随机变量的密度**

(1) 联合概率密度

1) 2)

(2) 分布函数：

(3) 边缘概率密度：

(4) 条件概率密度：

**4.常见二维随机变量的联合分布**

(1) 二维均匀分布： ,

(2) 二维正态分布：(

**5.随机变量的独立性和相关性**

和的相互独立::

（离散型） （连续型）

和的相关性：

相关系数时，称和不相关， 否则称和相关

**6.两个随机变量简单函数的概率分布**

离散型： 则：

连续型： 则：

，

**7.重要公式与结论**

(1) 边缘密度公式：

(2)

(3) 若服从二维正态分布 则有：

1)

2) 与相互独立，即与不相关。

3)

4) 关于Y=y的条件分布为：

5) 关于的条件分布为：

(4) 若与独立，且分别服从 则：

(5) 若与相互独立，和为连续函数， 则和也相互独立。

**随机变量的数字特征**

**1.数学期望**

离散型：；

连续型：

性质：

(1)

(2)

(3) 若X和Y独立，则 (4)

**2.方差**：

**3.标准差**：，

**4.离散型：**

**5.连续型：**

性质：

(1)

(2)与相互独立，则

(3)

(4) 一般有

(5)

(6)

**6.随机变量函数的数学期望**

(1) 对于函数

为离散型：；

为连续型：

(2) ;; ;

**7.协方差**

**8.相关系数**  ,阶原点矩 ; 阶中心矩

性质：

(1)

(2)

(3)

(4)

(5) ，其中

，其中

**9.重要公式与结论**

(1)

(2)

(3) 且 ，其中

，其中

(4) 下面5个条件互为充要条件：

注：与独立为上述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：来自总体的个相互独立且与总体同分布的随机变量，称为容量为的简单随机样本，简称样本。

统计量：设是来自总体的一个样本，）是样本的连续函数，且中不含任何未知参数，则称为统计量

样本均值：

样本方差：

样本矩：样本阶原点矩：

样本阶中心矩：

**2.分布**

分布：，其中相互独立，且同服从

分布： ，其中且， 相互独立。

F分布：，其中且，相互独立。

分位数：若则称为的分位数

**3.正态总体的常用样本分布**

(1) 设为来自正态总体的样本，

则：

1) 或者

2)

3)

4)

**4.重要公式与结论**

(1) 对于，有

(2) 对于，有；

(3) 对于，有

(4) 对于任意总体，有