### The mathematics behind programming

*Hanhui LI*

*Computer programming is an art form,* like the creation of poetry or music.

– Donald Knuth

目录

[第一部分 微分与级数 - 4 -](#_Toc272491934)

[费马引理 - 4 -](#_Toc1336799334)

[罗尔定理 - 4 -](#_Toc612491624)

[拉格朗日中值定理 - 4 -](#_Toc1257604497)

[柯西中值定理 - 5 -](#_Toc1024727305)

[洛必达(L’ Hospital)法则 - 5 -](#_Toc1920449842)

[泰勒(Taylor)中值定理1 - 5 -](#_Toc321280084)

[泰勒(Taylor)中值定理2 - 6 -](#_Toc980483230)

[函数曲线的凹凸性 - 6 -](#_Toc1339623179)

[无穷级数 - 7 -](#_Toc828214305)

[柯西审敛原理 - 7 -](#_Toc1956307928)

[幂级数 - 7 -](#_Toc1692710326)

[三角函数 - 8 -](#_Toc1666577273)

[傅里叶级数 - 8 -](#_Toc535019490)

[第二部分 线性代数 - 9 -](#_Toc558538441)

[矢量 - 9 -](#_Toc704556850)

[矩阵 - 9 -](#_Toc262148392)

[单位矩阵 - 9 -](#_Toc1439064347)

[加法运算律 - 9 -](#_Toc1393647515)

[数乘运算率 - 10 -](#_Toc429646776)

[三角矩阵 - 10 -](#_Toc1233343018)

[转置矩阵 - 10 -](#_Toc1283942682)

[共轭矩阵 - 11 -](#_Toc1308971318)

[共轭转置 - 11 -](#_Toc1058461758)

[矩阵乘法运算律 - 11 -](#_Toc1959718605)

[代数余子式 - 11 -](#_Toc1033900196)

[行列式 - 11 -](#_Toc1470406295)

[特征值与特征向量 - 12 -](#_Toc2024274036)

[迹 - 12 -](#_Toc1537787278)

[对称矩阵 - 12 -](#_Toc625089701)

[正定性 - 12 -](#_Toc392603583)

[Hermitian矩阵 - 13 -](#_Toc1418655897)

[正交矩阵 - 13 -](#_Toc1986211885)

[酉矩阵 - 13 -](#_Toc1777342227)

[Jacobian矩阵 - 13 -](#_Toc293279419)

[块矩阵 - 13 -](#_Toc672225268)

[Schur补 - 13 -](#_Toc178612409)

[正交三角分解 - 14 -](#_Toc1904103204)

[谱分解 - 14 -](#_Toc461242034)

[奇异值分解 - 14 -](#_Toc1826383415)

[Polar分解 - 15 -](#_Toc2042289334)

[满秩分解 - 15 -](#_Toc1525706537)

[LUP分解 - 15 -](#_Toc1595022179)

[cholesky分解 - 15 -](#_Toc499396952)

[第三部分 概率论 - 17 -](#_Toc998479788)

[基础定义 - 17 -](#_Toc1012579258)

[事件集合 - 17 -](#_Toc1759170378)

[概率函数 - 17 -](#_Toc1969174797)

[概率函数的基础属性 - 17 -](#_Toc1050329262)

[条件概率 - 18 -](#_Toc568328094)

[贝叶斯定理 - 18 -](#_Toc2030497649)

[无关性 - 19 -](#_Toc911352266)

[第四部分 数论 - 20 -](#_Toc1244164258)

[基础定义 - 20 -](#_Toc620413367)

[第五部分 图论 - 21 -](#_Toc1254352984)

[第六部分 数值优化 - 22 -](#_Toc63639489)

[第七部分 概率数据结构 - 24 -](#_Toc142035417)

[第八部分 量子计算与量子信息 - 25 -](#_Toc1334921702)

[参考 - 26 -](#_Toc1267385305)

### 微分与级数

#### 费马引理

设函数在点处的某邻域内有定义，并且在处可导，如果对任意的，有

 (或)，

那么

通常称导数等于零的点为函数的驻点(或稳定点，临界点)。

#### 罗尔定理

如果函数满足

(1)在闭区间上连续

(2)在开区间内可导

(3)在区间端点出的函数值相等，即

那么在内至少有一点，使得

#### 拉格朗日中值定理

如果函数满足

(1)在闭区间上连续

(2)在开区间内可导

那么在内至少有一点，使等式

 (0.1)

成立。

#### 柯西中值定理

如果函数及满足

(1)在闭区间上连续

(2)在开区间内可导

(3)对任一，****，区间端点出的函数值相等，即

那么在内至少有一点，使等式

 (0.2)

成立。

#### 洛必达(L’ Hospital)法则

1. 当时，函数**及都趋于0；**

(2)在点的某去心邻域内，**及**都存在且**；**

(3)存在（或为无穷大）,

则

 (0.3)

#### 泰勒(Taylor)中值定理1

如果函数在处具有阶导数，那么存在的一个邻域，对于该邻域内的任一，有

 (1.1)

其中

 (1.2)

*提示：依此对1.1式求导可得，在所考虑邻域内阶导数。，因此1.2得证（洛必达法则）。*

1.1式称为在处带有佩亚诺余项的阶泰勒公式。

#### 泰勒(Taylor)中值定理2

如果函数在的某个邻域内具有阶导数，那么对任一，有

 (1.3)

其中

 (1.4)

其中是在和其中之间的某个值。

*提示：依据柯西中值定理 , 是在和之间。继续可以得到，得证。*

1.3式称为带有拉格朗日余项的阶泰勒公式。

#### 函数曲线的凹凸性

定理 设函数在上连续，在内具有一阶和二阶导数，那么

1. 若在内，则在上的图形是凹的；
2. 若在内，则在上的图形是凸的。

#### 无穷级数

定义 如果级数的部分和数列有极限，即



那么称无穷级数收敛，这时极限叫做这级数的和，并写成

**

如果没有极限，那么称称无穷级数发散。

#### 柯西审敛原理

定理（柯西审敛原理） 级数收敛的充要条件为：对于任意给定的正数，总存在正整数，使得当时，对于任意的正整数，都有



成立。

#### 幂级数

如果给定一个定义在区间上的函数序列

，，，...，，...

那么由这函数序列构成的表达式



称为定义在区间上的函数项无穷级数，简称函数项级数。

函数项技术中最简单常见的一类级数即幂级数，他的形式是

 (1.x)

假设函数在点的某邻域内能展开成幂级数，则展开式叫做函数在点处的泰勒展开式。

 (1.y)

#### 三角函数

*三角函数和傅里叶级数的区别*

#### 傅里叶级数

定理 设周期为的周期函数满足收敛定理的条件，则他的傅里叶级数展开式为

 ， (1.z)

### 线性代数

#### 矢量

线性代数的核心是两个运算符，都和矢量相关，即加法运算符和乘法运算符。将两个矢量和相加得到，将将两个矢量和分别乘以和得到和，将结果相加得到线性组合。

点积（内积）指两个同样维度的矢量对应元素相乘的和。点积为零表示两个矢量相互垂直。

矢量长度：。

单位矢量：即长度为1的矢量。

Schwarz不等式：

三角不等式：

#### 矩阵

**定义：** 由个数排成的行列的数表称为行列的矩阵，简称矩阵。记为：



矢量是特殊的矩阵。

#### 单位矩阵

主对角上的元素都为1，即

#### 加法运算律





#### 数乘运算率









矩阵加减法和数乘称为矩阵的线性运算。

#### 三角矩阵

三角矩阵是方阵。它指对角线元素上或下的元素都为0。如果是上面的元素都为0，称之为上三角矩阵，反之成为下三角矩阵。

上三角矩阵

下三角矩阵

#### 转置矩阵

把矩阵的行和列相互交换产生的矩阵称为的转置矩阵，

矩阵的转置满足以下运算律：







*证明：*

*令，则，*

*所以*

#### 共轭矩阵

首先给出共轭复数的定义，即两个实部相等，虚部互为相反数的复数互为**共轭复数(conjugate complex number)**。

把矩阵的元素全部转换为共轭复数，转换的结果称为的共轭矩阵。

#### 共轭转置

求矩阵的共轭再转置称为的共轭专职，记为，或者。

#### 矩阵乘法运算律

结合律：

左分配率：

右分配律：

***矩阵乘法不满足交换律***。

#### 代数余子式

把的正方矩阵（称为阶矩阵）中元素所在的第行和第列划去，留下的阶行列式叫做元素的余子式，记作。将余子式再乘以的次幂称为元素的代数余子式。

#### 行列式

一个的正方矩阵的行列式记为或者。一个矩阵的行列式可表示如下，



一个的正方矩阵的行列式等于其任意行（或列）的元素与对应的代数余子式的乘积之和，即：



行列式为0的矩阵称之为奇异矩阵，反之称为非奇异矩阵。

#### 特征值与特征向量

阶矩阵的一个特征值和对应的特征向量是满足的标量以及非零向量。其中为特征向量，为特征值。

的所有特征值的全体，叫做的谱，记为。

#### 迹

阶矩阵的对角元素之和称为矩阵的迹（trace），记作，即



#### 对称矩阵

对称矩阵满足条件

#### 正定性

阶对称矩阵如果满足对所有非零向量，对应的二次型



若满足，称为正定矩阵。若则为负定矩阵，若则为半正定矩阵。若既非半正定，也非半负定，则为不定矩阵。矩阵是正定的当且仅当其特征值都是正数。

#### Hermitian矩阵

Hermitian矩阵满足条件。对称矩阵是特殊的Hermitian矩阵。也叫做self-adjoint矩阵。

#### 正交矩阵

正交矩阵满足条件。

#### 酉矩阵

酉矩阵满足条件。

#### Jacobian矩阵

Jacobian矩阵是函数的一阶偏导数以一定方式排列成的矩阵。

可表示成如下形式，



#### 块矩阵

矩阵作为一个数表，当遇到行数和列数较高的矩阵时，我们可以把整个矩阵分割成若干个小块，每个小块叫做矩阵的子块 ，以子块为元素的，形式上的矩阵叫做分块矩阵。

#### Schur补

对分块矩阵，采用消元法将其化简。

设分块矩阵为，则

称为Schur补。

#### 正交三角分解

需要进行正交化和三角化。

设，那么可唯一的分解为或，其中和为酉矩阵。和分别为正上三角矩阵和正下三角矩阵。

由于矩阵的列是线性无关的。利用Schmidt正交化和单位化方法，得到一组正交向量组再单位化得到一组标准正交向量组，

那么

#### 谱分解

休息休息

#### 奇异值分解

引理1:对于任何一个矩阵都有

引理2：对于任何一个矩阵都有与都是半正定的Hermite矩阵

设，是的特征值，是的特征值，它们都是实数。如果记

定理：设,那么，

同时，我们称0，为矩阵的正奇异值，简称奇异值。

定理：设，是矩阵的个 奇异值，那么存在阶酉矩阵和阶酉矩阵使得

，其中且满足

我们称此定理为奇异值分解定理。称表达式

为矩阵的奇异值分解式

由此可知矩阵的列向量就是的标准正交特征向量；而矩阵的列向量就是的标准正交特征向量。

#### Polar分解

设矩阵是可以作用于矢量空间的线性运算符。那么存在酉矩阵和算符和使得

其中矩阵和满足，。如果可逆，那么酉矩阵有唯一值。

#### 满秩分解

设为任意矩阵，则存在，使得。B称为列满秩矩阵，称为行满秩矩阵。

将矩阵A进行初等行变换，即将矩阵变换为最少行数不全为0元素的矩阵作为。选取其中线性无关的两列，进而选取对应原始矩阵的两列作为B，选取中不全为0的行作为。[参考](https://www.bilibili.com/video/BV1ci4y1m7FM/?spm_id_from=333.337.search-card.all.click)

#### LUP分解

休息休息

#### cholesky分解

把矩阵分解为如下形式，即称为cholesky分解。

或

1. 正定矩阵的对角元素都是正的。

2. 正定矩阵的Schur补也是正定矩阵。参考[证明](http://mlnotes.com/2013/08/22/cholesky.html)

### 概率论

#### 基础定义

研究概率学的起点是四个关键术语的定义。它们是：实验，样本结果，样本空间，和事件。后三个来自于经典集合论，第一个提供了将真实世界现象投射为概率学术语的概念上的机制。

实验要求其本身可以重复（理论上）无穷多次，并且具有规范定义的可能结果集。每个实验的可能结果称之为样本结果，所有的样本结果构成样本空间。使用表示属于。任何指定的样本结果集合，包括多个样本结果，整个样本空间或者空集构成一个事件。

#### 事件集合

定义在样本空间上的事件相关的操作称作集合代数。

定义1：

事件集合和事件集合的交集写做，表示事件结果同时属于和。

事件集合和事件集合的并集写做，表示事件结果属于或，或者同时属于和s。

定义2：

如果事件集合和事件集合的交集为空集，即，那么事件集合和事件集合互斥。

所有存在于属于样本空间，但不属于事件集合的事件组成的集合，叫做的补集。

#### 概率函数

如果包含有限个成员，那么概率函数满足

**公理1** 如果是定义于上的任意事件，那么

**公理2** 

**公理3** 和是定义在上的任意两个互斥的事件集合，那么

#### 概率函数的基础属性

属性1 

属性2 

属性3 如果那么

属性4 对任何时间，

属性5 对两两互斥的事件，，，，有

属性6 

属性7 对定义在上的任意事件，，，，有







#### 条件概率

设和是定义在上的任意两个事件集合，。的条件概率即是已经发生时的概率，写做。其值为





#### 贝叶斯定理

设是个事件的集合，每个事件发生的概率大于零，所有个事件的并集构成事件集合（即），且时，。对定义于上的任意事件,其概率。那么对任何有





#### 无关性

无关性定义，如果那么两个事件和是无关的。

### 数论

#### 基础定义

你好啊

### 图论

我不好

### 数值优化

设在物理世界中，有一种物理问题可以通过实验观测到它，并且可以在个变化条件下重复观测次并得到个观测值。个变化条件可以是一维空间（，如时间空间），三维空间（，如三维物理世界）、或者某个更高维度的抽象空间~~，我们可以将这个变化条件表示为矢量~~。

我们将进行的次观测实验得到的数据表示为，观测本身的目的是为了确定和矢量****之间的转换关系。而实际观测值和理论值之间存在着误差，所以我们将观测值表示为

 (1)

和矢量****之间的转换关系可能是线性关系，也可能是非线性关系，而且常常无法得到明确的解析式。如何和矢量****之间的转换关系由个参数确定，那么我们将参数表示为， (1)式可以写为

 (2)

其中称为误差。如果确定了最优的参数，那么也就确定了。当选取最优解时，以下函数取最小值，

 (3)

其中，称为损失函数。即要求选取最优解时，误差最小。

假设损失函数可微，那么损失函数可以使用泰勒展开表示，

 (4)

其中称为梯度，

 (5)

称为塞黑矩阵，

 (6)

### 概率数据结构

Etete

### 量子计算与量子信息

### 参考

1. 《高等数学第七版》, 同济大学。
2. Introduction to Linear Algebra, fifth version, Gilbert Strang, Massachusetts Institute of Technology.
3. An introduction to mathematical statistics and its applications, Sixth Edition, Richard J. Larsen and Morris L. Marx, Pearson.
4. <https://zhuanlan.zhihu.com/p/52890135，矩阵分解讲解。>
5. Quantum Computation and Quantum Information, 10th Anniversary Edition, Michael A. Nielsen & Isaac L. Chuang