代数、拓扑学、微分学和

优化理论

计算机科学与工程

Jean Gallier和Jocelyn Quaintance

计算机与信息科学系

宾夕法尼亚大学费城分校，PA 19104，美国，电子邮件：jean@cis.upenn.edu

C让·加利尔

2019年7月28日

二

# 目录

目录3

1. 引言17
2. 组、环和字段19
   1. 群，亚群，陪衬。……………………………….. 19
   2. 循环群。…………………………………………33
   3. 环和场。……………………………………….. 36

I线性代数43

1. 向量空间、基、线性映射45
   1. 向量空间。………………………………………….. 45
   2. 索引族；和符号pi∈i ai。……………………47
   3. 线性独立，子空间。……………………………52
   4. 向量空间的基。………………………………….. 57
   5. 矩阵。………………………………………………65
   6. 线性地图。………………………………………….. 69
   7. 商空间。……………………………………….. 77
   8. 线性形式和对偶空间。……………………………78
   9. 总结。…………………………………………….. 81
2. 矩阵和线性映射83
   1. 用矩阵表示线性图。……………………83
   2. 基矩阵的变化。………………………………….. 93
   3. 哈尔基向量和小波一瞥。…………………96
   4. 基的变化对矩阵的影响。……………………113个
   5. 总结。………………………………………………117个
3. 直接和119
   1. 和，直接和，直接积。…………………………119
   2. 秩零定理；格拉斯曼关系。………………128
   3. 总结。………………………………………………134个

三

1. 行列式135
   1. 排列，排列的签名。………………………135岁
   2. 交替多行地图。………………………………139
   3. 行列式的定义。…………………………………142个
   4. 逆矩阵和行列式。……………………………149个
   5. 线性方程和行列式系统。…………………152
   6. 线性映射的行列式。………………………………153
   7. 凯莱-汉密尔顿定理。………………………………154个
   8. 永久性。……………………………………………159
   9. 更多读数。…………………………………………161个
2. 高斯消元，Lu，Cholsky，梯队形式163
   1. 激励例子：曲线插值。………………………163个
   2. 高斯消去。……………………………………167
   3. 初等矩阵和行操作。………………………172个
   4. Lu因子分解。…………………………………………175个
   5. pa=lu因子分解。……………………………………181个
   6. 定理7.5~的证明。……………………………………189个
   7. 处理迂回错误；旋转策略。…………………195年
   8. 三对角矩阵的高斯消去。……………………196年
   9. SPD矩阵和Cholesky分解。…………………198
   10. 减少的行梯队形式。…………………………………207个
   11. 自由变量，齐次系统。……………………213个
   12. RREF的唯一性。………………………………………216个
   13. 用RREF求解线性系统。…………………………218
   14. 基本矩阵和列运算。…………………225
   15. 变换和扩张~。……………………………226
   16. 总结。………………………………………………231个
   17. 问题。………………………………………………233个
3. 向量范数和矩阵范数245
   1. 赋范向量空间。……………………………………245
   2. 矩阵规范。…………………………………………256
   3. 附属规范。………………………………………261
   4. 涉及从属规范的不平等。………………………268个
   5. 矩阵的条件数。………………………………270个
   6. 规范的应用：不一致线性系统。……………279
   7. 序列和序列的极限。………………………………280个
   8. 矩阵指数。……………………………………283个
   9. 总结。………………………………………………286个
   10. 问题。………………………………………………288个
4. 求解线性系统的迭代方法295

|  |  |
| --- | --- |
| 9.1向量和矩阵序列的收敛性。………………. | 二百九十五 |
| 9.2迭代法的收敛性。……………………………. | 二百九十八 |
| 9.3雅各比、高斯-赛德尔和放松的方法。………………… | 三百 |
| 9.4方法的收敛性。………………………………… | 三百零八 |
| 9.5三对角矩阵的收敛方法。…………………. | 三百一十一 |
| 9.6总结。……………………………………………… | 三百一十五 |
| 9.7问题。……………………………………………… | 三百一十六 |
| 对偶空间，对偶性 | 三百一十九 |
| 10.1双空间E和线性形式。………………………… | 三百一十九 |
| 10.2 E和E之间的配对和二元性。………………………. | 三百二十四 |
| 10.3对偶定理。……………………………………. | 三百二十九 |
| 10.4超平面和线性形式。……………………………… | 三百三十五 |
| 10.5线性映射和矩阵的转置。……………………. | 三百三十七 |
| 10.6四个基本子空间。…………………………… | 三百四十五 |
| 10.7总结。……………………………………………… | 三百四十八 |
| 11欧几里得空间 | 三百五十一 |
| 11.1内部产品，欧几里得空间。……………………………. | 三百五十一 |
| 11.2欧几里得空间的正交性和对偶性。…………………. | 三百六十 |
| 11.3线性图的伴随。…………………………………… | 三百六十七 |
| 11.4正交基底的存在和建造。………………. | 三百七十 |
| 11.5线性等轴测（正交变换）。…………………. | 三百七十七 |
| 11.6正交组，正交矩阵。…………………… | 三百八十 |
| 11.7罗德里格斯公式。…………………………………… | 三百八十二 |
| 11.8可逆矩阵的QR分解。……………………… | 三百八十五 |
| 11.9欧几里得几何的一些应用。……………………… | 三百九十 |
| 11.10总结。……………………………………………… | 三百九十一 |
| 11.11问题。……………………………………………… | 三百九十三 |
| 12任意矩阵的QR分解 | 四百零五 |
| 12.1正交反射。……………………………………. | 四百零五 |
| 12.2使用户主矩阵进行QR分解。…………………. | 四百一十 |
| 12.3总结。……………………………………………… | 四百二十 |
| 12.4问题。……………………………………………… | 四百二十 |
| 13个Hermitian空间 | 四百二十七 |
| 13.1赫米特空间，希尔伯特前空间。………………………… | 四百二十七 |
| 13.2线性映射的正交性、对偶性、伴随性。………………… | 四百三十六 |
| 13.3线性等轴测（也称为幺正变换）。…………… | 四百四十一 |
| 13.4一元群，一元矩阵。………………………… | 四百四十三 |
| 13.5厄米提反射和QR分解。…………………… | 四百四十六 |
| 13.6正交投影和对合。………………………. | 四百五十一 |
| 13.7双重规范。……………………………………………. | 四百五十四 |
| 13.8总结。……………………………………………… | 四百六十一 |
| 13.9问题。……………………………………………… | 四百六十二 |
| 14个特征向量和特征值 | 四百六十七 |
| 14.1线性图的特征向量和特征值。…………………… | 四百六十七 |
| 14.2还原为上三角形。…………………………. | 四百七十五 |
| 14.3特征值的位置。…………………………………… | 四百七十九 |
| 14.4特征值问题的处理。…………………………. | 四百八十二 |
| 14.5矩阵指数的特征值。………………………… | 四百八十五 |
| 14.6总结。……………………………………………… | 四百八十七 |
| 14.7问题。……………………………………………… | 四百八十八 |
| 15单位四元数和SO（3）中的旋转 | 四百九十九 |
| 15.1四元数的SU（2）组和斜场H。…………… | 四百九十九 |
| 15.2用su（2）中的四元数表示so（3）中的旋转。……… | 五百零一 |
| 15.3旋转RQ的矩阵表示。……………………… | 五百零六 |
| 15.4寻找代表旋转的四元数的算法。………. | 五百零八 |
| 15.5指数图exp:su（2）→su（2）。…………………….15.6四元数插值~。………………………………… | 五百一十一  五百一十三 |
| 15.7从SO（3）到SU（2）的“漂亮”部分不存在。……………. | 五百一十五 |
| 15.8总结。……………………………………………… | 五百一十七 |
| 15.9问题。……………………………………………… | 五百一十八 |
| 16谱定理 | 五百二十一 |
| 16.1引言。…………………………………………… | 五百二十一 |
| 16.2法向线性映射：特征值和特征向量。………………. | 五百二十一 |
| 16.3正态线性映射的谱定理。……………………… | 五百二十七 |
| 16.4自伴图和其他特殊线性图。……………………. | 五百三十二 |
| 16.5正态矩阵和其他特殊矩阵。…………………………… | 五百三十八 |
| 16.6瑞利-里兹定理和特征值交错。……………… | 五百四十一 |
| 16.7古兰-费希尔定理；摄动结果。……………… | 五百四十六 |
| 16.8总结。……………………………………………… | 五百四十九 |
| 16.9问题。……………………………………………… | 五百五十 |

## 17有限元法简介557

17.1一维问题：梁弯曲。…………………557个

17.2二维问题：弹性膜。………………568个

17.3时间相关边界问题。…………………………571

## 18个图形和拉普拉斯图形；基本事实579

18.1有向图、无向图、加权图。……………582个

18.2图的拉普拉斯矩阵。………………………………589

|  |  |
| --- | --- |
| 18.3图的正规拉普拉斯矩阵。……………………… | 五百九十三 |
| 18.4使用标准化切割进行图形聚类。……………………… | 五百九十七 |
| 18.5总结。……………………………………………… | 五百九十九 |
| 18.6问题。……………………………………………… | 六百 |
| 19光谱图绘制 | 六百零三 |
| 19.1图形绘制和能量最小化。……………………… | 六百零三 |
| 19.2图表示例。………………………………. | 六百零六 |
| 19.3总结。……………………………………………… | 六百一十 |
| 20奇异值分解与极性形式 | 六百一十三 |
| 20.1 f f的性质。………………………………………20.2平方矩阵的奇异值分解。………………. | 六百一十三  六百一十七 |
| 20.3方阵的极性形式。……………………………… | 六百二十 |
| 20.4矩形矩阵的奇异值分解。…………… | 六百二十三 |
| 20.5 KY Fan规范和Schatten规范。…………………………. | 六百二十六 |
| 20.6总结。……………………………………………… | 六百二十七 |
| 20.7问题。……………………………………………… | 六百二十七 |
| 21支持向量机和伪逆的应用 | 六百三十一 |
| 21.1最小二乘问题和伪逆问题。…………………. | 六百三十一 |
| 21.2伪逆函数的性质。……………………………. | 六百三十八 |
| 21.3数据压缩和SVD。………………………………… | 六百四十三 |
| 21.4主要成分分析（PCA）。………………………… | 六百四十五 |
| 21.5最佳仿射近似。………………………………… | 六百五十六 |
| 21.6总结。……………………………………………… | 六百五十九 |
| 21.7问题。……………………………………………… | 六百六十 |
| 22计算特征值和特征向量 | 六百六十三 |
| 22.1基本QR算法。…………………………………. | 六百六十五 |
| 22.2海森堡矩阵。……………………………………… | 六百七十一 |
| 22.3使用轮班使QR方法更有效。……………… | 六百七十七 |
| 22.4 Krylov子空间；Arnoldi迭代。…………………………. | 六百八十二 |
| 22.5克。………………………………………………. | 六百八十六 |
| 22.6赫米特案例；兰佐斯迭代。………………………… | 六百八十七 |
| 22.7动力方法。…………………………………………. | 六百八十八 |
| 22.8总结。……………………………………………… | 六百九十 |
| 22.9问题。……………………………………………… | 六百九十一 |
| 二仿射和射影几何 | 六百九十三 |
| 23仿射几何基础 | 六百九十五 |

|  |  |
| --- | --- |
| 23.1仿射空间。…………………………………………… | 六百九十五 |
| 23.2仿射空间示例。…………………………………. | 七百零四 |
| 23.3 Challes的身份。………………………………………. | 七百零五 |
| 23.4仿射组合，重心。……………………………. | 七百零六 |
| 23.5仿射子空间。………………………………………… | 七百一十一 |
| 23.6仿射独立性和仿射框架。………………………… | 七百一十七 |
| 23.7仿射图。……………………………………………. | 七百二十三 |
| 23.8仿射群。…………………………………………… | 七百三十 |
| 23.9仿射几何：一瞥。………………………………. | 七百三十二 |
| 23.10仿射超平面。………………………………………. | 七百三十六 |
| 23.11仿射空间的交集。………………………………… | 七百三十八 |
| 24在向量空间中嵌入仿射空间 | 七百四十一 |
| 24.1“帽结构”或均质化。……………………… | 七百四十一 |
| 24.2 E的仿射框架和E\_的底座。…………………………… | 七百四十八 |
| 24.3 E\_的另一种结构。………………………………… | 七百五十一 |
| 24.4将仿射映射扩展到线性映射。………………………… | 七百五十四 |
| 射影几何基础 | 七百五十九 |
| 25.1为什么是投影空间？……………………………………. | 七百五十九 |
| 25.2投影空间。………………………………………… | 七百六十四 |
| 25.3投影子空间。……………………………………… | 七百六十九 |
| 25.4投影框架。………………………………………. | 七百七十二 |
| 25.5投影图。………………………………………… | 七百八十六 |
| 25.6在两个投影帧之间找到一个同形。…………… | 七百九十二 |
| 25.7仿射补丁。…………………………………………. | 八百零五 |
| 25.8仿射空间的射影完成。……………………… | 八百零八 |
| 25.9充分利用无限远的超平面。…………………… | 八百一十三 |
| 25.10交叉比。………………………………………… | 八百一十六 |
| 25.11同系物和同系物的固定点。…………………. | 八百二十 |
| 25.12射影几何中的对偶性。……………………………… | 八百三十四 |
| 25.13超平面的交叉比。………………………………… | 八百三十八 |
| 25.14真实投影空间的复杂性。……………………. | 八百四十 |
| 25.15射影空间上的相似结构。……………………. | 八百四十二 |
| 25.16射影几何的一些应用。……………………… | 八百五十一 |
| 三双线性形式的几何 | 八百五十七 |
| 26卡坦-迪乌登定理 | 八百五十九 |
| 26.1线性等轴测的卡坦-迪乌顿定理。……………. | 八百五十九 |
| 26.2仿射等距图（刚性运动）。……………………………. | 八百七十一 |
| 26.3仿射图的不动点。………………………………… | 八百七十三 |
| 26.4仿射等距线和固定点。…………………………… | 八百七十五 |
| 26.5仿射等轴测的卡坦-迪乌顿定理。……………. | 八百八十一 |
| 赫米特空间27等距图 | 八百八十五 |
| 27.1卡坦-迪乌登定理，赫米特案例。………………… | 八百八十五 |
| 27.2仿射等距线（刚性运动）。……………………………. | 八百九十四 |
| 28双线性形式的几何；维特定理 | 八百九十九 |
| 28.1双线性形式。…………………………………………. | 八百九十九 |
| 28.2倍线性形式。………………………………………. | 九百零七 |
| 28.3正交性。…………………………………………… | 九百一十一 |
| 28.4线性图的伴随。…………………………………… | 九百一十六 |
| 28.5与倍线性形式相关的等轴测图。…………………… | 九百一十八 |
| 28.6完全各向同性子空间。………………………………… | 九百二十二 |
| 28.7维特分解。……………………………………… | 九百二十八 |
| 28.8辛群。………………………………………. | 九百三十六 |
| 28.9正交群和卡坦-迪乌顿定理。…………… | 九百四十 |
| 28.10维特定理。…………………………………………  四代数：pid's，ufd's，noetherian环，张量， | 九百四十七 |
| PID上的模块，正常形式 | 九百五十三 |
| 29多项式、理想和PID | 九百五十五 |
| 29.1多片式。……………………………………………… | 九百五十五 |
| 29.2多项式。……………………………………………. | 九百五十六 |
| 29.3欧氏多项式划分。…………………………… | 九百六十二 |
| 29.4理想、PID和最大公约数。…………………… | 九百六十四 |
| 29.5 K[X]中的因子分解和不可约因子。…………………… | 九百七十二 |
| 29.6多项式的根。……………………………………. | 九百七十六 |
| 29.7多项式插值（拉格朗日、牛顿、赫米特）。……………. | 九百八十三 |
| 30个湮灭多项式；一次分解 | 九百九十一 |
| 30.1湮灭多项式和最小多项式。……………. | 九百九十三 |
| 30.2对角化线性映射的最小多项式。……………… | 九百九十五 |
| 30.3线性地图的通勤族。…………………………. | 九百九十八 |
| 30.4主分解定理。…………………………一千零一  30.5约旦分解。……………………………………1007  30.6幂零线性映射和约旦形式。………………………1010 | |

30.7总结。………………………………………………一千零一十六

30.8问题。………………………………………………一千零一十七

## 31 UFD，诺特环，希尔伯特基定理1019

31.1唯一因子分解域（因子环）。…………………1019

31.2中国剩余定理。……………………………1033

31.3诺特环和希尔伯特基定理。…………………1039

31.4进一步的读数。…………………………………………一千零四十三

## 32张量代数1045

32.1线性代数预备：对偶空间和对。……………一千零四十七

32.2张量积。…………………………………………一千零五十二

32.3张量积的基。…………………………………1064

32.4张量积的一些有用同构。…………………1065

32.5张量积的对偶性。…………………………………一千零六十九

32.6张量代数。…………………………………………一千零七十五

32.7对称张量幂。…………………………………1082

32.8对称功率的基础。…………………………………一千零八十六

32.9对称幂的一些有用同构。…………………一千零八十九

32.10对称功率的对偶性。………………………………1089

32.11对称代数。………………………………………一千零九十三

32.12问题。………………………………………………一千零九十六

## 33外部张量幂和外部代数1099

33.1外张量幂。……………………………………一千零九十九

33.2外部权力基础。…………………………………1104

33.3一些对外部力量有用的同构。…………………1107

33.4外部力量的双重性。…………………………………一千一百零七

33.5外部代数。………………………………………1111

33.6 Hodge-操作员。……………………………………1115

33.7左右挂钩~。……………………………………一千一百一十九

33.8测试可分解性~。…………………………………一千一百二十九

33.9格拉斯曼PLU–克尔方程和格拉斯曼方程~。…………一千一百三十二

33.10向量值交替形式。……………………………1135

33.11问题。………………………………………………一千一百三十九

## 34模块介绍；PID 1141上的模块

34.1交换环上的模块。……………………………一千一百四十一

34.2模块的有限表示。………………………………1150 34.3交换环上模的张量积。……………1156

34.4 PID上的扭转模块；初级分解。………………一千一百五十九

34.5 PID上有限生成的模块。………………………1165

34.6鳞片环的延伸。……………………………1181

## 35正态形式；有理正态形式1187

35.1与自同态有关的扭转模块。…………1187

35.2有理规范形式。………………………………1195

35.3有理规范形式，第二版。……………………1202 35.4重新访问约旦表格。…………………………………一千二百零三

35.5史密斯标准形状。……………………………………一千二百零六

|  |  |
| --- | --- |
| V拓扑学、微分学 | 一千二百一十九 |
| 36拓扑 | 一千二百二十一 |

36.1度量空间和赋范向量空间。………………………一千二百二十一

36.2拓扑空间。………………………………………1228

36.3连续功能、限制。………………………………1237

36.4连接装置。…………………………………………1245

36.5紧凑型机组和局部紧凑型空间。……………………1254

36.6第二可数和可分空格。………………………1265

36.7连续压实度。……………………………………一千二百六十九

36.8完整的度量空间和紧凑性。………………………一千二百七十五

36.9公制空间的完成。………………………………1278

36.10收缩映射定理。…………………………1285

36.11连续线性和多线性地图。………………………1289

36.12赋范向量空间的完成。…………………………一千二百九十六

36.13赋范仿射空间。……………………………………1299

36.14进一步的阅读。…………………………………………一千二百九十九

1. 关于分形1301的绕道

37.1迭代函数系统和分形。………………………1301

## 38微积分1309

38.1方向导数、总导数。………………………一千三百零九

38.2雅可比矩阵。………………………………………1323

38.3隐函数和反函数定理。…………………1331 38.4切线空间和微分。……………………………1335

38.5二阶及更高阶导数。……………………1336

38.6泰勒公式，联邦航空局迪布鲁诺公式。………………………1341

38.7向量场，协变导数，方括号。…………………一千三百四十五

38.8进一步的读数。…………………………………………一千三百四十七

|  |  |
| --- | --- |
| 六、优化理论准备 | 一千三百四十九 |
| 39实值函数的极值 | 一千三百五十一 |

39.1局部极值和拉格朗日乘数。………………………1351

39.2使用二阶导数求极值。………………………一千三百六十一

39.3利用凸性求极值。……………………………一千三百六十四

39.4总结。………………………………………………一千三百七十四

## 40牛顿法及其推广1375

40.1牛顿的实变元实函数法。…………1375

40.2牛顿方法的推广。…………………………1376

40.3总结。………………………………………………一千三百八十二

## 41二次优化问题1383

41.1二次优化：正定情况。………………一千三百八十三

41.2二次优化：一般情况。……………………1392

41.3在单位球面上最大化二次函数。……………1397

41.4总结。………………………………………………一千四百零二

## 42舒尔补充和应用1403

42.1舒尔补充。………………………………………一千四百零三

42.2 SPD矩阵和舒尔互补。…………………………一千四百零六

42.3半定矩阵和舒尔补。………………1407

七、线性优化1409

## 43凸集，锥，H多面体1411

43.1什么是线性规划？ . ………………………………一千四百一十一

43.2仿射子集、凸集、超平面、半空间。……………1413

43.3锥体、多面体锥体和H多面体。……………………1416

## 44线性程序1423

44.1线性规划、可行解、最优解。……………一千四百二十三

44.2基本可行解和顶点。…………………………1429

## 45单纯形算法1437

45.1单纯形算法背后的想法。………………………1437

45.2单纯形算法。……………………………一千四百四十六

45.3如何有效地执行旋转步骤。…………………….1453 45.4使用tableaux的单纯形算法。………………………1457

45.5单纯形法的计算效率。…………………一千四百六十六

## 46线性规划与对偶1469

46.1法卡斯引理的变体。………………………………1469

46.2线性规划中的对偶定理。……………………一千四百七十四

46.3补充松弛条件。…………………………一千四百八十二

46.4标准形式线性程序的对偶性。…………………1484

46.5双单纯形算法。………………………………1487

46.6原始对偶算法。…………………………………一千四百九十二

|  |  |
| --- | --- |
| 八、非线性优化 | 一千五百零三 |
| 47希尔伯特空间基础 | 一千五百零五 |

47.1投影引理，对偶性。……………………………1505

47.2希尔伯特空间的Farkas–Minkowski Lemma。……………………1522

## 48优化理论的一般结果1525

48.1优化问题；基本术语。……………………1525

48.2优化问题解的存在性。………………1528

48.3二次函数的最小值。……………………………1533

48.4椭圆函数。………………………………………一千五百三十九

48.5无约束问题的迭代法。…………………1542

48.6无约束问题的梯度下降法。……………一千五百四十六

48.7变步长梯度下降收敛。……………一千五百五十一

48.8任意标准的最陡下降。………………………1556

48.9牛顿求最小值的方法。………………………一千五百五十八

48.10共轭梯度法；无约束问题。……………1562

48.11约束优化的梯度投影。………………1574

48.12约束优化的惩罚方法。…………………1576。

48.13总结。………………………………………………一千五百七十八

## 49非线性优化导论1581

49.1可行方向的锥体。………………………………一千五百八十一

49.2有效约束和合格约束。……………………一千五百八十八

49.3卡鲁什-库恩-塔克条件。…………………………一千五百九十四

49.4平等约束最小化。……………………………一千六百零六

49.5硬边支持向量机；第一版。…………………一千六百一十一

49.6硬边支持向量机；第二版。…………………一千六百一十五

49.7拉格朗日对偶和鞍点。…………………………一千六百二十四

49.8弱二元性和强二元性。…………………………………1633

49.9明确处理平等约束。………………………1641

49.10硬边支持向量机的对偶。………………1644

49.11共轭函数和勒让德对偶函数。…………………一千六百四十九

49.12获得更有用的双程序的一些技术。…………1659

49.13 Uzawa的方法。…………………………………………一千六百六十三

49.14总结。………………………………………………一千六百六十九

## 50次梯度和次微分1671

50.1扩展实值凸函数。……………………….1673 50.2次梯度和次微分。……………………………1682

50.3次梯度和次微分的基本性质。………………一千六百九十四

50.4次微分的附加属性。………………………1700。

50.5适当凸函数的最小值。……………………1704

50.6拉格朗日框架的推广。……………………一千七百一十

50.7总结。………………………………………………一千七百一十四

## 51双上升法；ADMM 1717

51.1双重上升。……………………………………………一千七百一十九

51.2增广拉格朗日和乘数法。……………1723

51.3 ADMM：乘法器的交替方向法。………………一千七百二十八

51.4行政管理的衔接。……………………………………一千七百三十一

51.5停止标准。…………………………………………一千七百四十

51.6 ADMM的一些应用。………………………………1741

51.7 ADMM在1-范数问题中的应用。……………………一千七百四十四

51.8总结。………………………………………………一千七百四十九

|  |  |
| --- | --- |
| 九、机器学习应用 | 一千七百五十一 |
| 52岭回归和套索回归 | 一千七百五十三 |

52.1岭回归。…………………………………………一千七百五十三

52.2套索回归（`1-正则化回归）。……………………一千七百六十三

52.3总结。………………………………………………一千七百六十九

## 53个正定核1771

53.1正定核的基本性质。……………………1771

53.2正核的希尔伯特空间表示。…………………一千七百八十二

53.3核PCA。……………………………………………一千七百八十六

53.4ν-sv回归。…………………………………………一千七百八十九

## 54软边界支持向量机1799

54.1软边界支持向量机；（SVMS1）。…………………1802

54.2软边界支持向量机（SVMS2）。…………………1812

54.3软边界支持向量机；（SVMS20）。…………………一千八百一十九

54.4软保证金SVM；（SVMS3）。…………………………………一千八百三十四

54.5软边界支持向量机；（SVMS4）。…………………1837

54.6软保证金SVM；（SVMS5）。…………………………………一千八百四十五

54.7支持向量机方法总结与比较。…………………一千八百四十八

|  |  |
| --- | --- |
| 十、附录 | 一千八百六十一 |
| 希尔伯特空间中的一个完全正交族 | 一千八百六十三 |

A.1总正交族，傅立叶系数。…………………1863

A.2希尔伯特空间2（k）和里兹费希尔定理。……………一千八百七十一

## b zorn的引理；一些应用1881

B.1佐恩引理的陈述。…………………………………1881 b.2矢量空间中基存在的证明。………………1882

B.3最大适理想的存在性。…………………………1883

参考文献1885

16目录

第一章

# 介绍

十七

18第1章。引言

第二章

# 组、环和字段

在接下来的四章中，我们回顾了基本的代数结构（群、环、域、向量空间），重点介绍了向量空间。综述了线性代数的基本概念，如向量空间、子空间、线性组合、线性独立性、基、商空间、线性映射、矩阵、基的变化、直和、线性形式、对偶空间、超平面、线性映射的转置。

## 2.1群体、亚群体、陪衬

实数集r有两个运算，即：r×在+andr→r（加法）和：r×r→r（乘法）下，形成一个交换群，满足使r

R−0=R转化为下的阿贝尔群。回想一下组的定义。

定义2.1.A群是一个具有二元运算·：G×G→G的集合G，它将元素A·B∈G与每对元素A、B∈G相关联，并具有以下性质：（W.R.T.·）。更明确地说，这意味着以下方程都是关联的，具有一个恒等元e∈g，并且GA，B，cis中的每个元素都是可逆的∈g：

（g1）a·（b·c）=（a·b）·c（结合性）；

（g2）a·e=e·a=a.（身份）；

（g3）对于每个a∈g，有一些a−1∈g，这样a·a−1=a−1·a=e.（逆）。群G是交换的（或交换的），如果

a·b=b·a，对于所有a，b∈g。

一个集合m加上一个操作·：m×m→m和一个只满足条件（g1）和（g2）的元素e被称为单体。例如，自然数的集合n=0,1，…，n，…是加法下的（交换）单体。然而，它不是一个群体。

下面给出了一些组的例子。

十九

例2.1。

1. 整数的集合z=…，−n，…，−1,0,1，…，n，…是加法下的交换群，具有标识元素0。然而，z=z−0不是乘法组。
2. 有理数（p，q∈z，q=0的分数p/q）6的集合q是一个加上的交换群，单位元为0。集合q=q−0也是一个带单位元素1的乘法下的交换群。
3. 对于任何非空集合s，f:s→s的双射集合，也称为s的排列，是函数组合下的一个组（即f和g的乘法是组合g\_f），同一元素是同一函数ID。只要S有两个以上的元素，这个群就不是阿贝尔群。集合s=1，…，n的置换群通常表示为sn，称为n元素上的对称群。
4. 对于任意正整数p∈n，在z上定义一个关系，表示为m n（mod p），如下所示：

m n（mod p）iff m−n=kp，对于某些k∈z。

读者将很容易地检查这是一个等价关系，而且它与加法和乘法兼容，这意味着如果m1 n1（mod p）和m2 n2（mod p），那么m1+m2 n1+n2（mod p）和m1m2 n1n2（mod p）。因此，我们可以定义等价类集（mod p）的加法运算和乘法运算：

[m]+[n]=[m+n]

和

[m]·[n]=[mn]。

读者将很容易地检查添加剩余类（mod p）是否诱导了一个[0]为零的阿贝尔群结构。该组表示z/pz。

1. 具有实（复）系数的n×n可逆矩阵集是矩阵乘法下的一个群，其单位元为单位矩阵。这个群称为一般线性群，通常用gl（n，r）（或gl（n，c））表示。
2. 具有实（或复）系数的n×n可逆矩阵a的集合，使得det（a）=1是矩阵乘法下的一个群，其中的单位元是单位矩阵。这个群称为特殊线性群，通常用sl（n，r）（或sl（n，c））表示。
3. 具有实数系数的n×n矩阵q的集合，这样

q q>=q>q=in

是矩阵乘法下的一个群，其中的单位元素是单位矩阵；我们有q−1=q>。这个群称为正交群，通常用O（n）表示。

1. 具有实数系数的n×n可逆矩阵q的集合，使得

q q>=q>q=in和det（q）=1

是矩阵乘法下的一个群，其中的单位元素是单位矩阵；如（6）所示，我们得到q−1=q>。这个群称为特殊正交群或旋转群，通常用so（n）表示。

（5）–（8）中的组是n≥2的非贝林族，但作为阿贝尔族（但O（2）不是阿贝尔族）的SO（2）除外。

通常用+表示阿贝尔群G的运算，在这种情况下，元素A∈G的逆A−1用−A表示。

组的标识元素是唯一的。事实上，我们可以证明一个更普遍的事实：

提案2.1.如果一个二元运算·：m×m→m是关联的，如果e0∈m是左恒等式，e00∈m是右恒等式，这意味着

e0·a=a表示所有a∈m（g2l）

和

a·e00=a，所有a∈m，（g2r）

则e0=e00。

证据。如果方程（g2l）中a=e00，我们得到

e0·e00=e00，

如果我们在方程（g2r）中让a=e0，我们得到

e0·e00=e0，

因此，e0=e0·e00=e00，

如要求。

命题2.1意味着一个单体的同一元素是唯一的，并且由于每个群都是一个单体，所以一个群的同一元素是唯一的。此外，一个组中的每个元素都有一个唯一的逆矩阵。这是一个更普遍的事实的结果：

提案2.2.在具有单位元e的单元体m中，如果某个元素a∈m有左逆a0∈m和右逆a0∈m，这意味着

a0·a=e（g3l）

和

a·a00=e，（g3r）

则a0=a00。

证据。利用（g3l）和e是一个身份元素的事实，我们有

（a0·a）·a00=e·a00=a00。

同样，使用（g3r）和e是一个标识元素的事实，我们有

a0·（a·a00）=a0·e=a0。

然而，由于m是单体的，运算·是关联的，所以

a0=a0·（a·a00）=（a0·a）·a00=a00，

如要求。

注：公理（g2）和（g3）可以通过只要求（g2r）（正确身份的存在）和（g3r）（每个元素都存在一个正确的反义词）（或（g2l）和（g3l））而稍微减弱。证明群公理（g2）和（g3）遵循（g2r）和（g3r）是一个很好的练习。

定义2.2.如果一个G组有有限个n个元素，我们就说G是一个n阶的组。如果G是无限的，我们就说G有无限阶。群的阶通常用g\_表示（如果g是有限的）。

对于G组，对于任意两个子集R，S G，我们让

r s=r·s r∈r，s∈s。

特别地，对于任何g∈g，如果r=g，我们写下

GS=G·S S∈S，

同样地，如果S=G，我们写

r g=r·g r∈r。

从现在开始，我们将去掉乘号，并将g1 g2写为g1·g2。

定义2.3.让G成为一个群体。对于任意g∈g，定义lg，由g左平移，由lg（a）=ga，对于所有a∈g，和rg，由g右平移，由rg（a）=ag，对于所有a∈g。

通常使用以下简单的事实。

提案2.3.给定G组，lg和rg的翻译是双射。

证据。我们为LG展示了这一点，证明RG是相似的。

如果lg（a）=lg（b），那么ga=gb，在左边乘以g−1，我们得到a=b，所以lg注入。对于任何b∈g，我们有lg（g−1b）=gg−1b=b，所以lg是可预测的。因此，lg是双目标的。

定义2.4.给定G群，G的子集H是G iff的子群。

1. G的单位元E也属于H（E∈H）；
2. 对于所有的h1，h2∈h，我们有h1h2∈h；
3. 对于所有的h∈h，我们有h−1∈h。

下列命题的证明留作练习。

提案2.4.给定一个群G，子集H G是G的一个子群，当h1，h2∈h，则h1h−2 1∈h。

如果G组是有限的，那么可以使用以下标准。

提案2.5.给定有限群G，子集H G是G iff（1）e∈H的子群；

（2）H在乘法下闭合。

证据。我们只需要证明定义2.4的条件（3）成立。对于任何a∈h，由于左平移la是双射的，它对h的限制是内射的，而由于h是有限的，它也是双射的。由于e∈h，存在唯一的b∈h，因此la（b）=ab=e。但是，如果a−1是g中a的倒数，我们也有la（a−1）=aa−1=e，并且通过la的注入性，我们得到a−1=b∈h。

例2.2.

1. 对于任意整数n∈z，集合

nz=nk k∈z

是组Z的子组。

1. 矩阵集

gl+（n，r）=a∈gl（n，r）det（a）>0

是组gl（n，r）的子组。

1. sl（n，r）组是gl（n，r）组的一个子组。
2. O（N）组是GL（N，R）组的一个子组。
3. 群SO（n）是群O（n）的一个子群，是群SL（n，r）的一个子群。
4. 不难证明每2×2旋转矩阵r∈so（2）都可以写成

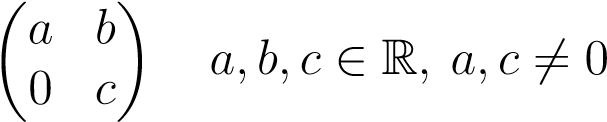
，0≤θ<2π。

然后，通过查看矩阵，可以将so（2）视为so（3）的一个子组。

作为矩阵

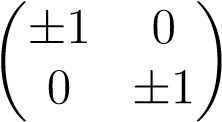
.

1. 形式的2×2上三角矩阵集



是组gl（2，r）的子组。

1. 由四个矩阵组成的集合V



是gl（2，r）组的一个子组，称为klein four组。

定义2.5.如果h是g的一个子群，g∈g是任意元素，那么gh形式的集合称为g中h的左陪集，hg形式的集合称为g中h的右陪集，左陪集（resp）。h的右陪集）诱导一个等价关系～定义如下：对于所有g1，g2∈g，

g1～g2 iff g1h=g2h

（响应g1～g2 iff hg1=hg2）。显然，是一个等价关系。

现在，我们声明以下事实：

提案2.6.给定G群和G的任意子群H，我们得到g1 h=g2h iff g2−1g1h=h iff g2−1g1∈h，对于所有g1，g2∈g。

证据。如果我们把双射同时应用于g1h和g2h，我们得到

并且，既然1∈h，我们得到g2−1g1∈h，反之，如果g2−1g1∈h，既然h是一个群，那么左边的翻译是h的双射，所以g2−1g1h=h，因此，g2−1g1h=h iff g2−1g1∈h。

由此可知，元素G∈G的等价类是coset-gh（resp。汞）。因为lg是h和gh之间的双射，所以cosets gh的基数都相同。图lg-1\_rg是左coset-gh和右coset-hg之间的双射，因此它们也具有相同的基数。由于不同的cosets-gh形成了g的一个划分，我们得到如下事实：

提案2.7.（拉格朗日）对于任何有限群G和G的任何子群H，H的阶H除以G的阶N。

定义2.6.给定有限群G和G的子群H，如果n=g和h=h，则比值n/h用（g:h）表示，称为g中h的指数。

指数（g:h）是g中h的左（右）陪集的个数。命题2.7可以表示为

| G=（G:H）H。

h在g中的左余割集（通常不是一个群）表示g/h。g/h的“点”是通过将coset中的所有元素“折叠”成单个元素来获得的。

例2.3.

1. 设n为任意正整数，并考虑z的nz子群（在加法下）。0的陪集是0，任意非零整数m∈z的陪集是

m+nz=m+nk k∈z。

用m除以n，我们得到一些独特r的m=nq+r，这样0≤r≤n-1。但是我们看到r是coset m+nz中最小的正元素。这意味着在z的nz子群的陪集和残基集0,1，…，n−1模n之间存在双射，或者与z/nz相等的双射。

1. gl（n，r）中sl（n，r）的陪集是矩阵集。

a sl（n，r）=a b b∈sl（n，r），a∈gl（n，r）。

由于a是可逆的，所以det（a）=06，如果det（a）>0，我们可以写a=（det（a））1/n（（det（a））-1/na，如果det（a）<0，我们可以写a=（det（a））1/n（（-det（a））-1/na。但如果det（a）>0，我们有（det（a））-1/na∈sl（n，r），如果det（a）>0，我们有−（−det（a））-1/na∈sl（n，r），如果det（a）<

0，所以coset asl（n，r）包含矩阵

（det（a））1/nin，如果det（a）>0，则−（−det（a））1/nin，如果det（a）<0。

由此可知，gl（n，r）和r中sl（n，r）的陪集之间存在双射。

1. gl+（n，r）中so（n）的陪集是矩阵集。

a so（n）=a q q∈so（n），a∈gl+（n，r）。

可以证明（使用矩阵的极性形式）gl+（n，r）中so（n）的陪集与n×n对称、正、定矩阵的陪集之间存在一个双射，它们是特征值严格为正的对称矩阵。

1. so（3）中so（2）的陪集是矩阵集。

q so（2）=q r r∈so（2），q∈so（3）。

群so（3）在r3中移动球面s2上的点，即对于任何x∈s2，

x 7→qx表示任意旋转q∈so（3）。

在这里，

s2=（x，y，z）∈r3 x2+y2+z2=1。

设n=（0,0,1）为球面s2上的北极。那么，不难证明so（2）正是so（3）的子群，使n保持不变。因此，陪集qso（2）中的所有旋转qr都映射到同一点qn∈s2，并且可以证明so（3）中的so（2）陪集与s2上的点之间存在双射。该图的可拓性与so（3）在s2上的作用是可传递的这一事实有关，也就是说，对于任何点x∈s2，都有一些旋转q∈so（3），使得qn=x。

在左陪集（或右陪集）上通过设置

（g1h）（g2h）=（g1g2）h，

但这种运算一般没有很好的定义，除非H子群具有特殊的性质。在示例2.3中，可以在（1）中定义coset的乘法，但在（2）和（3）中不可能。

允许在左陪集上定义乘法运算的子群H的性质是群同态核的典型性质，因此我们得出以下定义。

定义2.7.给定任意两组g和g0，函数\_：g→g0是同态iff

⑨（g1 g2）＝（g1）⑨（g2），对于所有g1、g2∈g。

取g1=g2=e（g），我们看到

⑨（e）=e0，

取g1=g和g2=g-1，我们可以看到

⑨（g−1）=（⑨（g））−1.

例2.4.

1. 所有m∈z中，由ω（m）=m mod n给出的ω：z→z/nz图是同态的。
2. 地图DET:GL（N，R）→同样，map det:r是同态，因为det（o（n）→r是同态。a b）=det（a）det（b）对于任意两个矩阵a，b

如果ω：g→g0和ψ：g0→g0是群同态，则ψ\_：g→g0也是同态。如果ω：g→g0是群的同态，如果h g，h0 g0是两个子群，则很容易检查

im h=\_（h）=\_（g）g∈h

是g0的子组，并且

⑨−1（h0）=g g（g）∈h0

是g的一个子群。特别是，当h0=e0\_时，我们得到了\_的内核ker\_。

定义2.8.如果ω：g→g g 0,0是群的同态，如果h g是g的子群，那么

im h=\_（h）=\_（g）g∈h，

用\_称为h的图像，G的子群，

ker\_ g∈g \_（g）=e0，

被称为\_的核。

例2.5。

1. 同态的核心是nz:z→z/nz。
2. 同态det的核：同态det的核：o（n）→rglis so（n，（rn））→r是sl（n，r）。类似地，核始终使用一个群同态的注入性的以下特征。

提案2.8.如果ω：g→g0是群的同态，那么ω：g→g0是注入式iff-ker\_ e\_。（我们还写了ker\_=（0）。）

证据。假设\_是注射的。既然\_（e）=e0，如果（g）=e0，那么（g）=（e），通过\_的注入率，我们必须有g=e，所以ker\_ e。

相反，假设Ker\_=\_e。如果直径（g1）=直径（g2），那么通过左侧乘以（直径（g1））-1，我们得到

e0=（\_（g1））-1（g1）=（（g1））-1（g2）、

并且，由于\_是同态（\_（g1））-1=\_（g1-1），所以

e0=（\_（g1））−1\_（g2）=（g1−1）\_（g2）=（g1−1g2）。

这表明g1−1g2∈ker，但由于ker\_e，我们得到g1−1g2=e，因此g2=g1，证明了\_是注射的。

定义2.9.我们说一个群同态，如果存在同态，那么它就是同态，因此

ψ\_=idg和\_ψ=idg0。（？）

如果ω是同构，我们就说G和G0组是同构的。当g0=g时，群同构称为自同构。

命题2.2证明中所用的推理表明，如果A群同态诳：g→g0是同态，则同态诳：g0→g满足条件（†）是唯一的。这种同态被称为\_−1。

左译lg和右译rg是g的自同构。

假设\_：g→g0是双射同态，让\_−1是\_的倒数（作为函数）。那么对于所有的a，b∈g，我们有

（−1（a）−1（b））=（−1（a））（−1（b））=ab，

因此，\_−1（a b）=−1（a）\_−1（b），

这就证明了磴-1是同态。因此，我们证明了以下事实。

提案2.9.双射群同态，\_：g→g0为同态。

观察该属性

gh=hg，对于所有g∈g.（）

等于右边的g−1乘以

g h g−1=h，对于所有g∈g，

上面的内容相当于

g h g−1 h，对于所有g∈g.（）

这是因为g h g−1 h意味着h g−1hg，而这对于所有g∈g。

提案2.10.设\_：g→g0为群同态。然后h=ker\_满足属性（），因此满足属性（）。

证据。我们有

⑨（g h g−1）＝（g）（h）（g−1）＝（g）e0（g）−1＝（g）（g）−1=e0，

对于所有的h∈h=ker\_和所有的g∈g，因此，根据h=ker\_的定义，我们得到了ghg−1 h。

定义2.10.对于G群，G的n子群是G iff的正规子群。

g n g−1=n，对于所有g∈g。

这用N c g表示。

命题2.10表明，同态的核ker a：g→g0是g的正规子群。

如果G是阿贝尔的，那么G的每个子群都是正常的。

考虑示例2.2。设r∈so（2）和a∈sl（2，r）为矩阵

.

然后

我们有

，

显然，ara−1∈/so（2）。因此（2）不是sl（2，r）的正规子群。同样的反例表明O（2）不是gl（2，r）的正规子群。设r∈so（2）和q∈so（3）为矩阵

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 然后 |  | . | | |  |  |  |
| 0 0 0 | 0 1 1 | |
| 我们有 |  |  |  |  |  |  |  |
| 1 qrq−1=0  零  0  ＝0  一 | 零  零  一  零  一  零 | 0 0−1 0 1 0  -1 1 0 0 0  0 0 0 1 0−1  -1  0℃。  零 |  | 01  零 |  | 零  零  1 |  |

观察qrq−1∈/so（2），so（2）不是so（3）的正规子群。设t和a∈gl（2，r）为下列矩阵

.

我们有

和

.

矩阵t是上三角形，但a t a−1不是，因此2×2上三角形矩阵组不是gl（2，r）的正规子群。

设q∈v和a∈gl（2，r）为下列矩阵

.

我们有

和

.

−−−

显然，aqa−1∈/v，这表明Klein四群不是gl（2，r）的正规子群。

读者应检查作为O（N，R）子群的nz、gl+（N，R）、sl（N，R）和so（N，R）子群是否为正常子群。

如果n是g的正规子群，由左陪集诱导的等价关系～与由右陪集诱导的等价关系相同（见定义2.5）。此外，这个等价关系是一个同余，这意味着：对所有人来说，

，然后，和

（2）如果g1n=g2n，则g1−1n=g2−1n。

因此，我们可以通过设置，在等价类模的集合g/上定义一个组结构。

（g1n）（g2n）=（g1g2）n.

定义2.11.设g为群，n为g的正规子群。通过定义（左）陪集的乘法得到的群

（g1 n）（g2n）=（g1 g2）n，g1，g2∈g

表示G/N，称G的商为N。元素G∈G的等价类Gn也表示G（或[G]）。图π：g→g/n由

π（g）=g=gn

是一个称为正则投影的群同态。

由于同态的核是正规子群，我们得到了以下非常有用的结果。

提案2.11.考虑到组\_：g→g0的同态，组g/ker\_和im\_=\_（g）是同态的。

证据。既然\_是对其图像的投射，我们可以假设\_是投射的，所以g0=im\_。我们定义了一个图：g/ker\_→g0，如下所示：

⑨（g）=⑨（g），g∈g.

我们需要检查这张地图的定义是否不依赖于coset g=g kerⅧ中选择的代表，并且它是同态的。如果g0是coset g ker\_中的另一个元素，这意味着对于某些h∈ker\_，g0=gH，那么

⑨（g0）=⑨（gH）=⑨（g）⑨（h）=⑨（g）e0=⑨（g）、

因为\_（h）=e0为h∈ker\_。这表明

⑨（g0）=（g0）=（g）=（g）、

因此，图\_定义良好。它是同态的，因为

图\_是注射剂，因为ker\_，iff g=e。图\_是注射剂，因为\_是注射剂。因此，如权利要求所述，Ⅷ是双射同态，因此是同态。

命题2.11被称为第一同构定理。

直接产品结构是构建群体的一种有效途径。

定义2.12.给定两组g和h，我们将g×h作为集合g和h的笛卡尔积，乘法运算·由下式给出

（g1，h1）·（g2，h2）=（g1g2，h1h2）。

立即证实G×H是G和H的直积。

同样，对于任意n组g1，…，gn，我们可以定义直接产物g1×·····×gn是类似的方式。

如果g是阿贝尔群，h1，…，hn是g的子群，情况就简单了。考虑一下地图

A:h1×····×hn→g

由a（h1，…，hn）=h1+····+hn给出，

用+来表示G群的运算，很容易验证A群是同态的，所以它的图像是G的一个子群，用h1+······+hn表示，称为H群的和。需要以下建议。

提案2.12。给定一个阿贝尔群G，如果h1和h2是g的任何子群，使得h1 h2=0，那么映射A是同构的。

A:h1×h2→h1+h2。

证据。根据定义，这张地图是投影的，所以我们只需要检查它是否是内射的。为此，我们证明了kera=（0,0）。我们有一个（a1，a2）=0 iff a1+a2=0 iff a1=−a2。由于a1∈h1和a2∈h2，我们看到a1，a2∈h1 h2=0，所以a1=a2=0，这证明了kera=（0,0）。

2.2。循环群

在命题2.12的条件下，即h1 h2=0，h1+h2组被称为h1和h2的直接和；它由h1 h2表示，我们有一个同构h1×h2～h1 h2。

## 2.2循环群

给定一个G群，单位为1，对于任意G∈G和任意自然数N∈N，定义Gn如下：

g0=1 gn+1=g·gn。

对于任何整数n∈z，我们定义gn为

G

G

n=（gn−1）（−n）如果nn<≥00。

以下特性易于验证：

g i·g j=gi+j（gi）−1=g−i gi·gj=gj·gi，

就我而言，J∈Z。

定义g的子集hgi by

hgi=gn n∈z。

下面的建议留作练习。

提案2.13.给定一个群G，对于任意元素G∈G，集Hgi是包含G的G的最小阿贝尔子群。

定义2.13.群G是循环的，如果有一个元素G∈G，那么g=hgi。具有这个性质的元素G∈G称为G的生成器。

实施例2.2中的Klein四组V是Abelian，但不是循环的。这是因为v有四个元素，但所有不同于同一性的元素都有2阶。

循环群是z的商，对于这个，我们使用z的一个基本性质。回想一下，对于任何n∈z，我们让nz表示n的倍数集，nz=n k k∈z。

提案2.14.对于某些n∈n，z的每个子群h的形式为h=nz。

证据。如果h是平凡组0并且两者之一，那么让nm=0或。如果−mh是非平凡的，对于任何非零元素都是正的，那么让nhbe是包含n的最小元素。

m∈h，我们也有h。根据命题2.13，−m∈h nz是中正整数的最小子群。

对于m 6=0的任何m∈h，我们可以写

m=nq+r，0≤r<n。

现在，从nz h开始，我们得到nq h，从m h开始，我们得到r=m nq∈h。但是，0≤r<n，与n的最小值相矛盾，所以r=0，h=nz。

由\_给出任何循环群（m）=gm。由于g generatesg，对于任何generatorg，该映射是可预测的。G的映射，我们可以定义一个映射，显然是一个：z→G组同态，所以让H=Ker\_作为它的核心。通过前面的观察，对于一些n∈z，h=nz，因此通过第一个同态定理，我们得到了一个同态，即：z/nz−→g。

从商群z/nz到g。显然，如果g有有限阶，那么g=n。总之，我们得到以下结果。

提案2.15。对于某些自然数n>0，每个循环群g要么与z同构，要么与z/nz同构。在第一种情况下，我们说g是无限循环群，在第二种情况下，我们说g是n阶循环群。

商群z/nz由余割m+nz=m+nk k∈z组成，其中m∈z，即在等价关系下z的等价类的定义如下：

x y iff x−y∈nz iff x y（mod n）。

我们也用x来表示x的等价类x+nz，或者如果我们想用[x]n更精确地表示，则可以用

x+y=x+y。

对于每一个x∈zx，在x的x类中都有一个唯一的代表，这样x mod n≤（x mod n的非负余数≤n−1）。为此

原因是，我们经常将z/nz与集合0，…，确保-1。更严格地说，我们可以通过定义+n给出0，…，n一个群结构。

x+n y=（x+y）mod n.

那么，很容易看出z0.，…，n−1与操作+n是一个具有与z/n同构的标识元素0的群。

2.2。循环群

我们还可以定义z/nz上的乘法运算·如下：

A·B=AB=AB模式N。

然后，很容易检查·············································这是一个变化莫测的戒指。我们通常抑制点并写a b而不是a·b.z/nz

变成一个

提案2.16。给定任意整数n≥gcd（1，对于any a，n）=1a.∈z，余数类a∈z/nz对于乘法iff是可逆的。

证据。如果a在z/nz中有倒数b，那么ab=1，这意味着

AB（N型）

对于某些k∈z，即ab=1+nk，这是贝佐特恒等式。

ab−nk=1

意味着gcd（a，n）=1。相反，如果gcd（a，n）=1，那么根据贝佐特的恒等式，存在u，v∈z，这样au+nv=1，

所以au=1−nv，也就是说，

Au 1（n型）

这意味着au=1，所以a在z/nz中是可逆的。

定义2.14.环z/nz可逆元素的群（乘法下）用（z/nz）表示。注意，该组是阿贝尔的，仅在n≥2时定义。

在群理论（z/nz）中，欧拉函数起着重要作用。

函数定义2.15.）的定义是这样的：给定任意正整数，\_（n）是整数的个数，n≥1，Eulera，其中1函数≤a（或Euler≤n，它相对地是n的素数；也就是说，gcd（a，n）=1.1

然后，根据2.16号提案，我们看到该集团（z/nz）有第\_（n）号订单。

n=4，我们haven=2，（（zz//24zz））==11，3，平凡的群体。为}。两个群都同构于群1，如果n是素数，根据命题2.16，我们看到n=3，（z/3z）=−1,12，1和for。自从

（

## 2.3环和场

群z、q、r、c、z/nz和mn（r）比阿贝尔群多，它们也是交换环。此外，q、r和c是字段。我们现在介绍环和场。

additiondefinition 2.16.）和：a a×ringa→是一个具有以下属性的seta（被称为乘法，用两个操作+填充）：a×a→a:（被称为

（r1）a是阿贝尔群w.r.t.+；

（r2）是关联的，并且具有一个标识元素1∈a；

（r3）是分布的w.r.t.+。

加法的单位元表示为0，a∈a的加法逆表示为−aa。更明确地说，一个环的公理是下列方程：对于所有a，b，c∈

|  |  |
| --- | --- |
| A+（B+C）=（A+B）+C（正的结合性） | （2.1） |
| A+B=B+A（交换性为+） | （2.2） |
| A+0=0+A=A（零） | （2.3） |
| A+（−A）=（−A）+A=0（加性逆） | 网络错误 |
| 网络错误 | 网络错误 |
| 网络错误 | 网络错误 |
| 网络错误 | 网络错误 |
| 网络错误  网络错误  网络错误  网络错误 | 网络错误 |
| 网络错误 | 网络错误 |
| 网络错误 | 网络错误 |

注意，（2.9）意味着如果1=0，那么a=0代表所有a∈a，因此a=0。环A=0 a is称为双元素环A，B的B。1∈A的环通常表示为=06 abis called。非平凡的。乘法

例2.6.

1. 加群Z、Q、R、C是交换环。
2. 对于任意正整数n∈n，z/nz组是一个加上的组。我们也可以定义一个乘法运算

a·b=ab=ab模式n，

对于安拉零安达，b∈1zas乘法单位。由此产生的环用表示。读者可以很容易地通过z/nz.20检查环公理是否满足要求。

1. 一个实系数变量中多项式的R[X]群是多项式乘法下的环。它是一个交换环。
2. 设为任意正整数。如果d不可被m形式的任何整数除，m∈n和m≥2，那么我们就说，8，9，12不是无平方的。IFD是无平方的。例如，d是任意平方自由整数，如果d=1,2,3,5,6,7,10是平方自由整数，但4d≥2，则实数集

Z[√D]=A+B√D∈R A，B∈Z

是一个交换环。如果z=a+b√d∈z[√d]，我们写z=a−b√d。注意，zz=a2−db2。

1. 同样，如果d≥1是一个正的无平方整数，那么复数的集合

Z[√−D]=A+Ib√D∈C A，B∈Z

Zis是一个交换环。如果z=a2+db2。其中z=a+d ib=1√是一个著名的例子，由∈z[√−d]研究，我们写z=a−ib√d。注意

高斯和Z[√−1]，也表示Z[i]，称为高斯整数环。

1. n×n矩阵群mn（r）是矩阵乘法下的一个环。然而，它不是交换环。
2. f连续函数f:（a，b）→r的C（a，b）组是操作g下的一个环，定义如下：

（f·g）（x）=f（x）g（x）

对于所有x∈（a，b）。

定义2.17.给定一个环a，对于任何元素a∈a，如果有元素的话。环A是积分b∈A

域（或除0之外没有零除数的。

这样b=0，对于allb 6=0a，本泰尔环和∈a。换句话说，一个积分域是一个非平凡交换环ab=0），如果，那么我们说，0 6=1，a是交换的，anda是零除数ab=0意味着a=0或例2.7。

1. 环Z，Q，R，C是积分域。
2. 一个实系数变量的多项式环r[x]是一个积分域。
3. 对于任意正整数，n∈n，我们有环z/nz。如果n是复合的，那么这个环就有零除数。例如，如果n=4，那么我们有2.2\_0（mod 4）。

读者应该证明z/nz是一个积分域，iff n是素数（使用命题2.16）。

1. 如果d是一个无平方正整数，如果d≥2，则环z[√d]是一个积分域。√−类似地，如果d 1是一个无平方正整数，则环z[d]是一个积分域。找到这些环的可逆元素是一个非常有趣的问题。
2. n×n矩阵的环mn（r）具有零除数。

环之间的同态是一个保持加法和乘法（以及0和1）的映射。

定义2.18。给定两个环a和b，a和b之间的同态是一个函数h:a→b，满足所有x，y∈a的下列条件：

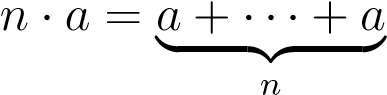
h（x+y）=h（x）+h（y）h（xy）=h（x）h（y）h（0）=0 h（1）=1。

实际上，因为b是一个加上的群，所以h（0）=0从

h（x+y）=h（x）+h（y）。

例2.8。

1. 如果a是一个环，对于任意整数n∈z，对于任意a∈a，我们将n·a定义为



如果n≥0（0·a=0），且

N A=（N）A

如果n<0.然后，地图

h（n）=n·1a

是一个环同态（其中1a是a的乘法恒等式）。

1. 给定任意实数λ∈r，评价图ηλ：r[x]→r定义为

ηλ（f（x））=f（λ）

对于每一个多项式f（x）∈r[x]都是一个环同态。

定义2.19.环自身称为非同态G：一个环同态M→A自同态G。f h=：idaa→and b fis an g=同构midb。一个来自aiff的同构有一个环

在群同构的情况下，同态G是唯一的，用H−1表示，很容易证明双射环同态H:A→B是同构。

定义2.20。给定环A，a的子集a0是a的子环，如果a0是a的子组（在加法下），则在乘法下闭合，并且包含1。

例如，我们有以下序列，其中入口符号左侧的每个环都是包含符号右侧的环的子环：

Z Q R C.

环√Z是Z[√D]和Z[√−D]的子环，环Z[√D]是R的子环，并且

环Z[−D]是C的子环。

如果h:a→b是环的同态，那么对于任何子环a0，图像h（a0）都是b的子环，对于任何子环b0，逆图像h−1（b0）是a的子环。

对于群，环同态H:A→B的核定义为

KER H=A∈A H（A）=0。

就像在群的情况下，我们有一个环同态的内射性的下列标准。证明与分组证明相同。

IFF建议2.17.KER H=0。（我们还写了，如果h:a→bkeris是环的同态，那么h=（0）。）h:a→b是内射的。

环同态的核是加群A的交换子群，但一般来说它不是A的子环，因为它可能不包含乘法恒等元1。但是，在乘法运算中，它满足以下闭包属性：

AB∈KER H和BA∈KER H代表所有A∈KER H和所有B∈A。

这是因为如果h（a）=0，那么对于所有b∈a，我们有h（ab）=h（a）h（b）=0h（b）=0和h（ba）=h（b）h（a）=h（b）0=0。

定义2.21。给定一个环A，满足以下性质的A的加性子群I

a b∈i和b a∈i表示所有a∈i和所有b∈a（理想）

被称为双面理想。如果A是一个交换环，我们简单地说是一个理想。

结果表明，对于任意环A和任意双面理想I，加性陪集A+I（带a∈a）的集A/I是一个称为商环的环。然后我们有下面的2.11命题的类似物，也叫做第一同构定理。

提案2.18。给定环H:A→B的同态，环A/KER H和im H=H（A）是同态的。

场是交换环k，其中a−0是乘法下的群。

定义2.22.如果集合k是一个环，则它是一个字段，并且以下属性保持不变：

（f1）0=16；

（F2）k=k−0是一个组w.r.t.（即，每个a=06都有一个与w.r.t.相反的组）；

（f3）是交换的。

如果不是交换的，但（f1）和（f2）保持不变，我们就说我们有一个斜场（或非交换场）。

注意，我们假设一个场的运算是交换的。这一惯例并不是普遍采用的，但由于对我们将遇到的大多数领域都是可交换的，所以我们也可以在定义中包括这一条件。

例2.9.

1. 环q、r和c是字段。
2. 多项式f（x），g（x）∈r[x]的（形式）分数集f（x）/g（x），其中g（x）不是零多项式，是一个场。
3. 连续函数f:（a，b）→r的环c（a，b），使得f（x）=06对于所有x∈（a，b）都是一个场。
4. 利用2.16号命题，很容易看出环z/pz是场iff p是素数。
5. 如果d是一个无平方正整数，如果d≥2，集合

是一个字段。如果b√d，那么很容易检查，如果z=06，那么z−1=z/（zz）。

1. 同样，如果d≥1是一个无平方的正整数，则复数的集合

q（√−d）=a+ib√d∈c a，b∈q

是一个字段。如果z=a+ib√d∈q（√−d）且z=a−ib√d，则很容易检查，如果z=06，则z−1=z/（zz）。

定义2.23。两个场k1和k2之间的同态h:k1→k2只是环k1和k2之间的同态。

然而，由于和是乘法下的群，域的同态必须是内射的。

证据。首先，观察任何x=06，

1=h（1）=h（x x−1）=h（x）h（x−1）

和

1=H（1）=H（X−1X）=H（X−1）H（X）、

所以h（x）=06和h（x−1）=h（x）−1。

但是，如果h（x）=0，我们必须得到x=0。因此，H是注射剂。

定义2.24.场同态H:k1→k2是同态，如果存在同态G:k2→k1，则G F=idk1，F G=idk2。从场到场的同构称为自同构。

然后，就像在环的情况下一样，G是唯一的，用H−1表示，双射场同态H:k1→k2是同态。

定义2.25。由于两个场之间的每一个同态h:k1→k2都是内射的，所以k1的图像f（k1）是k2的一个子场。我们说k2是k1的延伸。

例如，r是q的扩展，c是r的扩展。字段q（√d）和q（√−d）是√−q的扩展，字段r是q（√d）的扩展，字段c是

Q（d）的扩展。

定义2.26.如果系数为k的每一个多项式p（x）在k中都有根，则称k为代数闭域；也就是说，在k中有一些a∈k，使得p（a）=0。

可以看出，每个域k都有一些极小的扩张Ω，它是代数闭合的，称为k的代数闭包。例如，c是r的代数闭包。q的代数闭包称为代数数域。这个域由所有系数为q的多项式的零复数组成。

定义2.27。给定一个场k和一个自同构h:k→k的k，很容易检查集合

fix（h）=a∈k h（a）=a

由h固定的k元素的子字段是k的子字段，称为h固定的字段。

例如，如果d≥2为无平方，则图c:q（√d）→q（√d）由

C（A+B√D）=A−B√D

是q（√d）和fix（c）=q的自同构。

如果k是一个场，我们得到了由h（n）=n·1给出的环同态h:z→k。如果h是内射的，那么k包含z的副本，因为它是一个字段，所以它包含q的副本。在这种情况下，我们称k具有特征0。如果h不是内射的，那么h（z）是k的一个子环，因此是一个积分域，h的核是z的一个子群，根据命题2.14，对于某些p≥1，它必须是pz的形式。根据第一个同构定理，当p≥1时，h（z）与z/pz同构。但是，p必须是素数，因为z/pz是一个积分域iff，它是一个场iff，p是素数。素数p被称为k的特征，我们也说k是有限特征。

定义2.28。如果k是一个字段，那么

1. n·16=0，对于所有整数n≥1，在这种情况下，我们称k具有特征0，或
2. 有一些最小的素数p，使得p·1=0称为k的特征，我们称k为有限特征。

特征值为0的字段k包含q的副本，因此是无限的。正如我们将在第7.10节中看到的，有限域具有非零特性p。然而，存在非零特性的无限域。

第一部分

线性代数

四十三

第三章

# 向量空间、基、线性映射

## 3.1向量空间

对于每n≥1，设rn为n-元组x=（x1，…，xn）的集合。添加可扩展到RN，如下所示：

（x1，…，xn）+（y1，…，yn）=（x1+y1，…，xn+yn）。

我们还可以定义一个操作·：r×rn→rn如下：λ·（x1，…，xn）=（λx1，…，λxn）。

由此得到的代数结构具有一些有趣的性质，即向量空间的性质。向量空间定义如下。

定义3.1.给定一个字段k（加+和乘），向量空间超过

k（或k-向量空间）是一个集合e（向量的）加上两个运算+：e×e→e（称为向量加成），和·：k×e→e（称为标量乘），满足以下条件：所有α，β∈k和所有u，v∈e；

（v0）e是一个Abel群w.r.t.+，单位元素为0；

（v1）α·（u+v）=（α·u）+（α·v）；

（v2）（α+β）·u=（α·u）+（β·u）；

（v3）（αβ）·u=α·（β·u）；

（v4）1·u=u。

在（v3）中，表示K字段中的乘法。

四十五

给定α∈K和V∈E，元素α·V也用αV表示，K场常称为标量场。

除非另有规定，或者除非我们处理几个不同的字段，在本章的其余部分中，我们假设所有k向量空间都是相对于固定字段k定义的。因此，我们将k向量空间简单地称为向量空间。在大多数情况下，字段k是reals的字段r。

从（v0），向量空间总是包含空向量0，因此是非空的。（从（v1），我们得到−α）·v=−（α·v），α·0=0，α·（−v）=−（α·v）。从（v2），我们得到0·v=0，并且

公理的另一个重要结论是：对于任何一个λ∈k，如果λ6=0和λ·u=0，则u=0.u∈e和

实际上，由于λ6=0，它有一个乘法逆λ−1，因此从λ·u=0，我们得到

λ−1·（λ·u）=λ−1·0。

然而，我们只是观察到λ−1·0=0，从（v3）和（v4），我们已经

λ−1·（λ·u）=（λ−1λ）·u=1·u=u，

我们推断u=0。

备注：您可能会怀疑是否真的需要AXIOM（v4）。它能从其他公理推导出来吗？答案是否定的。例如，可以取e=rn并定义

·R×RN→RN by

λ·（x1，…，xn）=（0，…，0）

对于所有（x1，…，xn）∈rn和所有λ∈r，公理（v0）–（v3）都满足，但（v4）失败。使用尚未定义的基的概念可以给出一些不那么简单的例子。

字段k本身可以看作是一个向量空间，向量的加法是字段中的加法，标量的乘法是字段中的乘法。

例3.1。

1. r和c是r上的向量空间。
2. Rn和Cn是R上的向量空间，Cn是C上的向量空间。
3. 多项式的环r[x]是r上的向量空间，c[x]是r和c上的向量空间，n×n矩阵的环mn（r）是r上的向量空间。
4. 连续函数f:]a，b[→r的环c（]a，b[）是r上的向量空间。

设e为向量空间。我们要定义线性组合和线性独立的重要概念。

在定义这些概念之前，我们需要讨论一个战略选择，它取决于如何解决，在处理诸如线性组合和线性依赖（或独立性）等概念时可能会减少或增加负担。这个问题与使用向量集和向量序列有关。

## 3.2索引族；和符号pi∈i ai

我们的经验告诉我们，最好使用向量序列；更好的是，索引向量族。（我们不是唯一一个选择序列超过集合的人，而且我们有很好的伙伴关系，例如，Artin[7]、Axler[10]和Lang[106]使用序列。然而，一些著名的作者，如LAX[110]使用套装。我们把它留给读者对这个问题进行调查。）

在给定集合A的情况下，回想一个序列是一个有序的n-元组（a1，…，an）∈a中元素的a，对于某个自然数n。序列的元素不必是不同的，顺序是重要的。例如，（a1，a2，a1）和（a2，a1，a1）是a3中两个不同的序列。它们的基础集是a1，a2。

我们刚刚定义的是有限序列，也可以看作是从1,2，…，n到集合A的函数；序列的第i个元素（a1，…，an）是函数下i的图像。这个观点是卓有成效的，因为它允许我们将（可数）无限序列定义为函数s:n→a。但是，为什么要将我们自己限制为有序集，如1，…，n或n作为索引集？

索引集的主要作用是唯一地标记每个元素，标记的顺序虽然方便，但并不重要。因此，很自然地将一个i-索引族的元素定义为函数a:i→a，其中i是任何被视为索引集的集合。因为函数A是由它的图决定的

（i，a（i））i∈i，

A族可以看作是成对的A=（I，A（I））I∈I。为了表示简单，我们写a i而不是a（i），并用（ai）i∈i∈i表示家族a=（i，a（i））i∈i，例如，如果i=r，g，b，y和a=n，对的集合

A=（R，2），（G，3），（B，2），（Y，11）

是一个索引族。元素2在族中出现两次，分别带有两个不同的标记r和b。

当索引集i完全有序时，一个族（ai）i∈i通常称为i序列。有趣的是，集合可以看作是家庭的特殊情况。实际上，一个集合a可以被看作是与同一函数对应的a-索引族（a，a）a∈i。

备注：索引族不应与多重集混淆。对于任何集合A，多集都与集合相似，只是集合的元素可能出现多次。例如，如果a=a、b、c、d，则a、a、b、c、c、d、d是多集。每个元素都以一定的多重性出现，但元素的顺序并不重要。例如，a具有多重性3。形式上，多集是一个函数s:a→n，或等价于一组对（a，i）a∈a。因此，多重集是n个元素的a-索引族，而不是n-索引族，因为不同的元素可能具有相同的多重性（如上例中的c和d）。索引族是序列的泛化，而多重集是集合的泛化。

我们还需要注意一个恼人的技术性，即定义形式pi∈i a i的和，其中i是任何有限的索引集，（ai）i∈i是某个集合a中的一个元素族，该集合a装备有一个二元运算+：a×a→a，它是结合的（公理（g1））和交换的。当我们定义线性组合时，就会出现这种情况。

问题是，二元运算+只告诉我们如何计算a的两个元素的a1+a2，但它不告诉我们三个以上元素之和是什么。例如，应该如何定义A1+A2+A3？

我们要做的是用一系列步骤来定义a1+a2+a3，每个步骤都涉及两个元素，有两种可能的方法可以做到这一点：a1+（a2+a3）和（a1+a2）+a3。如果我们的操作+不是关联的，那么它们是不同的值。如果它是关联的，那么a1+（a2+a3）=（a1+a2）+a3，但是指数1、2、3还有六个可能的排列，如果+不是交换的，这些值通常是不同的。如果我们的操作是交换的，那么所有六个置换都有相同的值。因此，如果+是关联的和交换的，那么可以直观地看出形式pi∈i a i的和不依赖于用于计算它的操作的顺序。

确实如此，但严格的证据需要归纳，令人惊讶的是，这样的证据牵涉其中。读者可以不加证明地接受这样一个事实，即形式pi∈i ai的和确实定义得很好，并直接跳到定义3.2。对于那些想看到血淋淋细节的人，我们开始吧。

首先，我们定义和pi∈i a i，其中i是不同自然数的有限序列，比如i=（i1，…，im）。如果i=（i1，…，i m）m≥2，我们用i−i1表示序列（i2，…，im）。我们根据I的尺寸M进行归纳，让

.

例如，如果我=（1,2,3,4），我们有

.

如果操作+不是关联的，则术语的分组很重要。例如，一般来说

A1+（A2+（A3+A4））=（6 A1+A2）+（A3+A4）。

但是，如果操作+是关联的，那么和pi∈i ai不应该依赖于i中元素的分组，只要它们的顺序保持不变。例如，如果i=（1,2,3,4,5），

j1=（1,2）和j2=（3,4,5），我们期望

.

这是事实，因为我们有以下的主张。

提案3.1.给定任意一个非空集a，装备有一个结合二元运算+：a×a→a，对于任意非空有限序列i的不同自然数，以及对于i到p的任何分区，非空序列i k1，…，i kp，对于某些非空序列k=（k1，…，kp），这样ki<kj意味着α<β对于所有α∈iki和所有β∈ikj，对于a中元素的每个序列（ai）i∈i，我们有

.

证据。我们根据i的大小n进行归纳。

如果n=1，那么我们必须有p=1和ik1=i，所以这个命题无关紧要。

接下来，假设n>1。如果p=1，那么ik1=i，公式很简单，所以假设p=2，然后写j=（k2，…，kp）。有两种情况。

案例1。序列ik1有一个单一元素，比如β，它是i的第一个元素。在这种情况下，通过删除i的第一个元素β，为从i获得的序列写c。通过

定义，

和

.

由于c=n−1，根据诱导假设，我们有

，

这就产生了我们的身份。

案例2。序列ik1至少有两个元素。在这种情况下，让β是i的第一个元素（因此是i k1的第一个元素），让I0是从i中通过删除第一个元素β获得的序列，让I0是从ik1中通过删除第一个元素β获得的序列，让I=2，…，p。回想一下j=（k2，…，kp）和k=（k1，…，kp）。序列I0有n-1个元素，因此通过应用诱导假设，我们得到

.

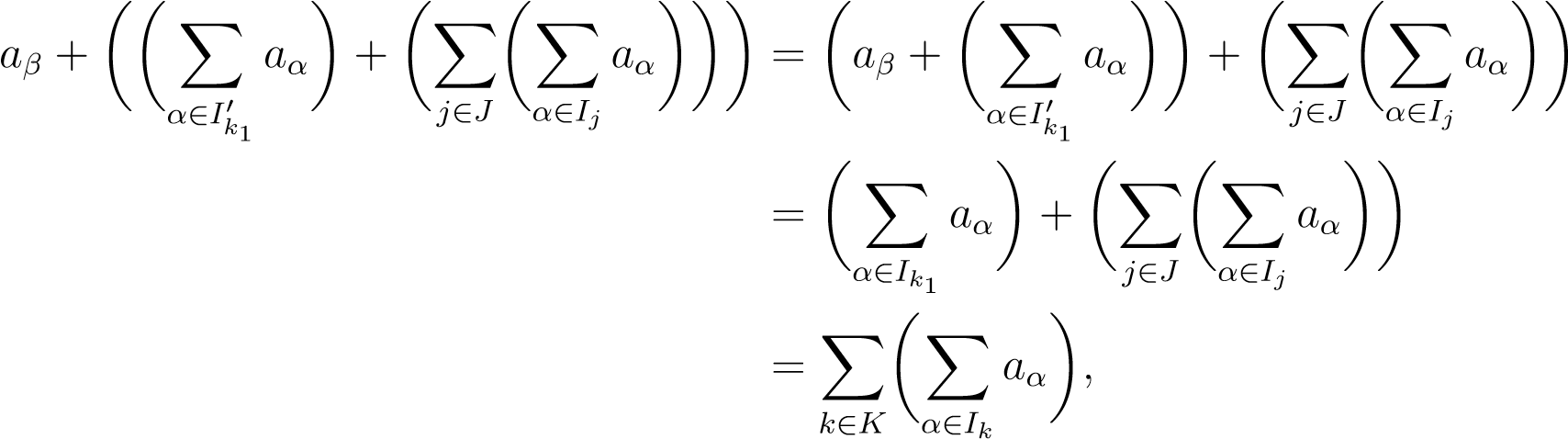
如果我们把左手边加在Aβ上，根据定义我们得到

X

α。

我一世

如果我们用关联性和一个索引和的定义把右手边加到β上，我们得到



如要求。

如果i=（1，…，n），我们也写而不是pi∈i ai。由于+是关联的，命题3.1表明总和独立于其元素的分组，这证明了符号a1+·····+an（没有任何括号）的使用是合理的。

如果我们也假设A上的结合二元运算是交换的，那么我们可以证明和pi∈i ai不依赖于索引集i的次序。

提案3.2.对于任意两个不同自然数的非空有限序列i和j，如果任意一个非空集a配备有结合和交换二元运算+：a×a→a，那么j是i的置换（换句话说，i和j的非空集是相同的），对于每个序列（ai）。A中元素的i∈i，我们有

十倍

aα=aα。α∈iα∈j

证据。我们通过归纳i中元素的数量p来进行，如果p=1，我们就得到i=j，这个命题就无关紧要了。

如果p>1，为了简化表示法，假设i=（1，…，p）并且j是i的置换（i1，…，ip）。首先，假设2≤i1≤p−1，让j0是通过删除i1从j获得的序列，i0是通过删除i1从i获得的序列，让p=（1,2，…，i1−1）和q=（i1+1，…，p-1，p）。观察到序列I0是序列P和Q的串联。通过应用于j0和i0的诱导假设，然后通过应用于i0及其分区（p，q）的命题3.1，我们得到

.

如果我们将左手边添加到ai1，根据定义，我们得到

X

α。J·J

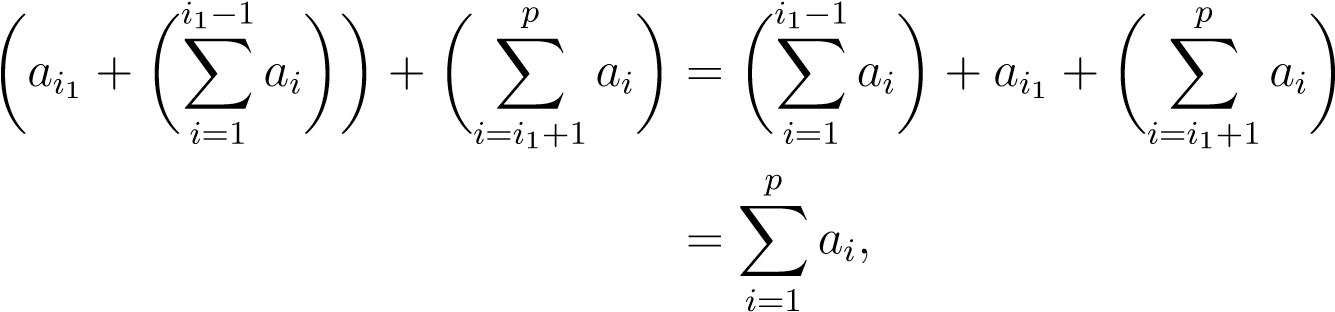
如果我们把右手边加在ai1上，我们得到

.

利用关联性，我们得到

，

然后使用结合性和交换性数次（更严格地说，在I1-1上使用归纳法），我们得到



如要求。

I1=1或I1=P的情况同样处理，但处理方式更简单，因为P=（或Q=）（其中（）表示空序列）。

做了所有这些之后，我们现在可以理解形式pi∈i a i的和，对于a中的任何有限索引集i和任何族a=（ai）i∈i，其中a是一个具有结合性和交换性的二元运算+的集。

事实上，由于i是有限的，它与集合1，…，n是双射的，对于一些n∈n，并且任何全序都对应一个置换（在这里我们用它的图像识别一个置换）。对于所有订单，我们都会定义

.

那么，对于其他的全部订货，我们有

，

既然我和是1，…，n的不同排列，根据命题3.2，我们有

.

因此，sum不依赖于i的总排序，我们将和pi∈i ai定义为所有总排序的公共值。

## 3.3线性独立性，子空间

向量空间的一个最有用的性质是有基。这意味着在每一个向量空间中，e，都有一组向量，e1，…，en，这样每一个向量，v∈e，都可以写成一个线性组合，

V=λ1e1+····+λnen，

对于ei，对于某些标量，λ1，…，λn∈k。此外，n-元组（λ1，…，λn），如上所述是唯一的。

当e有一个有限的基础，e1，…，en时，这种描述是很好的，但情况并非总是如此！例如，实多项式的向量空间r[x]没有有限基，而是有无限基，即

1，x，x2，…，xn，…

人们可能会想，一个向量空间是否可能有不同大小的基，甚至可能有一个有限的基和一个无限的基。稍后我们将看到这是不可能的；一个向量空间的所有基都有相同数量的元素（基数），这称为空间的维数。然而，我们有以下问题：如果一个向量空间有一个无限的基，e1，e2，…，，我们如何定义线性组合？我们允许线性组合吗

λ1e1+λ2e2+···

3.3。线性独立性，子空间

有无穷多非零系数？

如果我们允许具有无穷多非零系数的线性组合，那么我们必须理解这些和，并且只有当我们将这些和定义为具有s1=λ1e1和

sn+1=sn+λn+1en+1。

但是，我们如何定义这些限制呢？好吧，我们必须通过一个规范、一个度量或其他机制，在我们的空间上定义一些拓扑。这确实可以做到，这就是Banach空间和Hilbert空间的全部内容，但这似乎需要大量的机器。

避免极限的一种方法是将我们的注意力限制在只涉及有限多个向量的线性组合上。我们可能有无穷多的向量，但我们只形成包含有限多个非零系数的线性组合。从技术上讲，这可以通过引入有限支撑族来实现。这使我们能够操作由某个固定的无限集索引的标量族，但仍将这些族视为有限的。

考虑到这些动机，给定一个集合a，回想一下，i-索引族（ai）i∈i的元素（简而言之，一个族）是一个函数a:i→a，或等价于一组对（i，ai）i∈i。我们同意，当i=？，（ai）i∈i=？。族（a i）i∈i是有限的，如果我是有限的。

备注：在考虑家族（a i）i∈i时，没有理由假定我是有序的。关键的一点是，族中的每个元素都是由i元素唯一索引的。因此，除非另有规定，否则我们不假定索引集的元素是有序的。

如果a是恒等式为0的阿贝尔群（通常，当a是一个环或向量空间时），我们称一个族（a i）i∈i具有有限支持，如果ai=0表示所有i∈i−j，其中j是i（族的支持）的有限子集。

给定两个不相交集i和j，两族（ui）i∈i和（vj）j∈j的并集，表示为

（ui）i （v j）j（vj）j（i（j）定义如下：wk=Uk，如果k（wi）i；（ui）i；（vj）j j）定义如下：wk（wk）k天在k/∈i.给定一个族（ui）i∈i时，（ui）i∈i的一个子族是一个族（uj）j∈j，其中j是i的任何子集。

在本章中，除非另有规定，否则假定所有scalar族都有有限的支撑。

定义3.2.设e为向量空间。向量v∈e是e元素的族（ui）i∈i的线性组合，如果k中存在标量的族（λi）i∈i，则

.

当我=时，我们规定v=0。（根据命题3.2，形式pi∈iλi ui的和被很好地定义）我们说，如果对于k中的每个族（λi）i∈i，那么一个族（ui）i∈i是线性独立的，

X

λiui=0表示所有i∈i的λi=0。

我爱我

等价地，如果K中有标量的某个族（λi）i∈i，则该族（ui）i∈i是线性相关的，因此

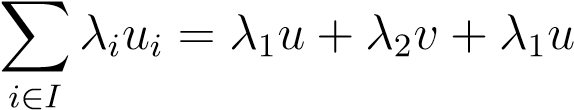
X

一些j∈i的λiui=0和λj=06。

我爱我

我们同意当我为∅时，族∅是线性独立的。

注意，定义向量族的线性组合而不是向量集的线性组合的优点是被组合的向量不必是不同的。例如，对于i=1,2,3和族（u，v，u）和（λ1，λ2，λ1），线性组合



有道理。在线性组合的定义中使用一组向量不允许这样的线性组合；这太限制了。

解开定义3.2，一个族（ui）i∈i是线性相关的，如果该族中的一些uj可以表示为该族中其他向量的线性组合。事实上，K中有一个标量家族（λi）i∈i，这样

X

一些j∈i的λiui=0和λj=06，

我想我

这意味着

.

注意，定义向量族而不是向量集的线性依赖性的原因之一是，我们的定义允许一个向量多次出现。这很重要，因为矩阵可能包含相同的列，我们想说这些列是线性相关的。集合的线性依赖性的定义不允许我们这样做。

上述结果还表明，一个族（ui）i∈i是线性独立的iff，或者i=∅，或者i由单个元素i和ui=06组成，或者i≥2，并且族中没有向量uj可以表示为族中其他向量的线性组合。

当i非空时，如果族（ui）i∈i是线性独立的，注意ui=06表示所有i∈i。否则，如果ui=0表示某些i∈i，则我们通过选取任何非零的λi并让λk=0表示所有k∈i，得到一个非平凡的线性依赖性pi∈iλiui=0，因为

3.3。线性独立性，子空间

λi0=0.如果i≥2，我们也必须有ui=6 uj代表所有i，j∈i代表i=6 j，因为否则我们通过选取任何非零λ的λi=λ和λj=−λ得到一个非平凡的线性依赖性，并且让λk=0代表所有k∈i代表k=6 i，j。

因此，线性独立的定义意味着一个非平凡的线性独立族实际上是一个集合。这解释了为什么某些作者选择定义向量集的线性独立性。这种方法的问题是线性依赖，即线性独立性的逻辑否定，然后只定义为一组向量。然而，正如我们前面所指出的，对于允许同一向量多次出现的族，定义线性相关性是非常可取的。

例3.2。

1. 任何两个不同的标量λ，μ=06在k中都是线性相关的。
2. 在r3中，向量（1,0,0）、（0,1,0）和（0,0,1）是线性无关的。
3. 在R4中，向量（1,1,1,1）、（0,1,1,1）、（0,0,1,1）和（0,0,0,1）是线性无关的。
4. 在r2中，向量u=（1,1），v=（0,1）和w=（2,3）是线性相关的，因为

W=2U+V。

一个族（i的j）即使当ni∈i是线性独立的iff（i=∅）时，每个有限子注。实际上，当nupj）j i∈jiλ为线性无关时，k中标量的族（λi）i∈i具有有限的支持，因此pλiui=0实际上意味着p j∈jλjuj=0 i∈i

对于i的有限子集j，当i是有限的时，我们通常假设它是集i=1,2，…，n。在这种情况下，我们将族（ui）i∈i表示为（u1，…，un）。

向量空间的子空间的概念定义如下。

定义3.3.在给定向量空间e的情况下，如果f是非空的，则e的子集f是e的线性子空间（或子空间），对于所有u，v∈f，以及所有λ，μ∈k，则λu+μv∈f。

很容易看出，e的子空间f确实是一个向量空间，因为·：e×e→e到f×f的约束实际上是一个函数·：f×f→f，而·：k×e→e到k×f的约束实际上是一个函数·：k×f→f。

很容易看出，子空间的任何交叉点都是子空间。因为f是非空的，如果我们选取任意向量u∈f，如果我们让λ=礹=0，那么λu+礹U=0u+0u=0，那么每个子空间都包含向量0。对于任意一个非空的有限指数集i，通过对i的基数的归纳，可以看出，如果（ui）i∈i是向量的任意族ui∈f，（λi）i∈i是标量的任意族，那么pi∈iλiui∈f。

子空间0将用（0）表示，甚至用0表示（略带滥用符号）。

例3.3.

1. 在r2中，向量u=（x，y）的集合如下：

X+Y=0

是子空间。

1. 在r3中，一组向量u=（x，y，z），使得

X+Y+Z=0

是子空间。

1. 对于任意n≥0，次数的多项式f（x）∈r[x]集至多n是r[x]的子空间。
2. 上三角n×n矩阵集是n×n矩阵空间的一个子空间。

提案3.3.对于任何向量空间e，如果s是e的任何非空子集，则包含s的e的最小子空间hsi（或跨度）是s元素的所有（有限）线性组合的集合。

证据。我们证明了S元素的所有线性组合的集跨度是E的一个子空间，作为练习，验证了包含S的每个子空间也包含跨度。

首先，跨度（pμsjv）j是非空的，因为它包含跨度（s）（非空）中的任意两个线性组合。如果），对于任意两个标量，u=pλ，μi∈i∈λirui，

且v=j∈j

λu+μv=λxλiui+μxμjvj

i∈i j∈j

=xλλiui+x祄祄jvj

i∈i j∈j

=xλi ui+x（λi+祄祄）ui+x祄祄祄jvj，

i∈i−j i∈i j∈j−i

它是指数集i j的线性组合，因此λu+μv∈SPAN（s），证明SPAN（s）是一个子空间。

有人可能会想，如果我们在组成线性组合的系数中添加额外的条件，会发生什么。这里有三个很重要的自然限制（和往常一样，我们假设我们的索引集是有限的）：

1. 考虑组合pi∈iλiui，其中

X

λi=1.

我爱我

这些被称为仿射组合。我们应该认识到每一个线性组合

pin II∈，如果我们把λiui看作仿射组合。例如，如果j=i k，uk=0，且λk=1−pi∈iλi，那么pj∈jkλ是一个指数，notjuj是一个仿射组合，并且

十倍

λiui=λjuj.

i∈i j∈j

然而，我们得到了新的空间。例如，在r3中，三个向量e1=（1,0,0）、e2=（0,1,0）和e3=（0,0,1）的所有仿射组合的集合是通过这三个点的平面。因为它不包含0=（0,0,0），所以它不是线性子空间。

1. 考虑组合pi∈iλiui，其中

λi≥0，对于所有i∈i。

这些被称为正（或圆锥）组合。事实证明，向量族的正组合是圆锥。它们在凸优化中表现得很自然。（3）考虑我们需要（1）和（2）的组合pi∈iλiui，即

，且所有i∈i的λi≥0。

这些被称为凸组合。对于任何有限的向量族，这些向量的所有凸组合的集合都是凸多面体。凸多面体在凸优化中起着非常重要的作用。

## 3.4向量空间的基

给定一个向量空间e，给定一个族（v i）i∈i，由零向量0和（vi）i∈i的所有线性组合组成的e的子集v很容易被看作是e的一个子空间。如果不是多余的话。具有这样一个“高效”生成族（称为基）的子空间起着重要作用，并激发了以下定义。

定义3.4.在给定向量空间e和e的子空间v的情况下，向量v i∈v的族（vi）i∈i跨越v或生成v，如果对于每一个v∈v，在k中有一些标量的族（λi）i∈i，这样

.

我们还说，（v i）i∈i的元素是v的生成元，v的范围是（vi）i∈i，或者由（vi）i∈i生成。如果e的子空间v是由有限族（vi）i∈i生成的，我们说v是有限生成的。跨越V且线性无关的族（ui）i∈i称为V的基。

例3.4.

1. 在r3中，向量（1,0,0）、（0,1,0）和（0,0,1）构成基。
2. 矢量（1，1，1，1），（1，1，−1，−1），（1，−1，0，0），（0，0，1，−1）构成了称为HAAR基的R4的基。这一基础及其对维数2n的推广在小波理论中是至关重要的。
3. 在次数最多为n的r[x]多项式的子空间中，多项式1、x、x2、…、xn构成一个基。
4. 伯恩斯坦多项式也构成了这个空间的基础。这些多项式在样条曲线理论中起着重要作用。

线性代数的第一个关键结果，即每个向量空间e都有一个基。我们从一个重要的引理开始，它使逐步构建基础的机制正式化。

引理3.4。给定一个向量空间E元素的线性独立族（ui）i∈i，如果v∈e不是（ui）i∈i的线性组合，则将v添加到族（ui）i∈i中得到的族（ui）i∈i∮k（v）是线性独立的（其中k/∈i）。

证明：对于任何族（λi）i∈ivof scalars in−p（祏k−1），μ具有一个逆（因为μv+pki∈iis a场），因此我们得到λiui=0。如果λi）uμi，显示=06，则

我爱我

V是（ui）i∈i的线性组合，与假设相矛盾。因此，μ=0。但是，我们有π∈iλi ui=0，由于族（ui）i∈i是线性独立的，所以我们有λi=0表示所有i∈i。

下一个定理一般成立，但对于没有有限生成集的向量空间，证明更为复杂。因此，在本章中，我们只证明有限生成向量空间的定理。

定理3.5。给定任意有限族s=（ui）i∈i生成向量空间e和任意线性无关子族l=（uj）j∈j（其中j i），e有一个基b，使得l b s。

证据。考虑线性无关族B的集合，使L B S。由于该集合是非空的和有限的，它有一些极大的元素，（即，具有最大基数的H I的S的子族B=（uh）H∈H），说B=（uh）H∈H。我们声称B生成E。实际上，如果B不生成E。当e为e时，则存在一些上∈s，它不是b中向量的线性组合（因为s生成e），带有p/∈h，然后，通过引理3.4，族b0=（uh）h∈h p是线性独立的，由于l b b0 s，这与b的最大性相矛盾，因此，b是e的基础，这样t帽子L B S.

注：定理3.5也适用于非有限生成的向量空间。在这种情况下，问题是要保证最大线性独立族B的存在，这样L B S。这种最大族的存在可以用Zorn引理表示，见附录B和此处给出的参考文献。

需要定理3.5的完全一般性的情况是向量的情况。

系数q上的空间r。数字1和√2是线性无关的。

在q上，所以根据定理3.5，线性独立族L=（1，√2）可以扩展到r的基b。因为r不可数，q可数，所以这样的基必须不可数！

基的概念也可以用极大线性独立族和极小生成族的概念来定义。

定义3.5.假设（vi）i∈i是向量空间e中的向量族，我们说（vi）i∈i是e的最大线性独立族，如果它是线性独立的，并且如果对于任何向量w∈e，通过将w添加到族（vi）i∈i得到的族（vi）i∈i∮k w是线性相关的。我们认为（vi）i∈i是e的极小生成族，如果它跨越e，并且对于任何指数p∈i，通过从族（vi）i∈i中删除vp得到的族（vi）i∈i−p不跨越e。

下面给出了表征基的有用性质的命题是引理3.4的直接结果。

提案3.6.给定一个向量空间e，对于e的向量的任何族b=（vi）i∈i，下列性质是等价的：

1. B是E的基础。
2. B是E的最大线性独立族。
3. B是E的最小生成族。

证据。假设（1）。因为B是基，所以它是一个线性独立的族。我们认为B是一个最大线性独立族。如果b不是最大线性无关族，则存在一个向量w∈e，使得通过将w加到b得到的b0族是线性无关的。然而，由于b是e的基础，向量w可以表示为b中向量的线性组合，这与b0是线性无关的事实相矛盾。

相反，假设（2）。我们认为b跨e，如果b不跨e，那么在b中有一个向量w∈e，它不是向量的线性组合，根据引理3.4，把w加到b得到的b0族是线性无关的。由于b是b0的一个合适的亚科，这与b是最大线性独立族的假设相矛盾。因此，b必须跨越e，因为b也是线性无关的，所以它是e的基础。

再次假设（1）。因为B是基，所以它是E的生成族，我们认为B是最小生成族。如果b不是最小生成族，那么b的一个合适的子族b0跨越e，那么每个w∈b−b0都可以表示为b0的向量的线性组合，这与b是线性独立的这一事实相矛盾。

相反，假设（3）。我们声称b是线性无关的。如果b不是线性独立的，那么一些向量w∈b可以表示为b0=b−w中向量的线性组合。由于b产生e，b0族也产生e，但b0是b的一个适当的亚科，与b的最小值相矛盾，因为b跨越e，是线性独立的，所以它是e的基础。

线性代数的第二个关键结果是，对于向量空间e的任意两个基（ui）i∈i和（vj）j∈j，索引集i和j具有相同的基数。特别地，如果e有n个元素的有限基，e的每个基都有n个元素，整数n称为向量空间e的维数。

为了证明第二个关键结果，我们可以使用下面的因steinitz而替换的引理。这一结果显示了向量空间有限线性独立族和有限生成族之间的关系。我们从引理的一个版本开始，它有点非正式，但比命题3.8中给出的更精确和更正式的公式更容易理解。技术上的困难与一些指数需要重新命名的事实有关。

提案3.7.（替换引理，版本1）给定向量空间e，让（u1，…，um）是e中的任何有限线性独立族，让（v1，…，vn）是任何有限族，这样每个ui都是（v1，…，vn）的线性组合。然后，我们必须使m≤n，并且用（u1，…，um）替换向量vj的m，这样在重命名vj的一些索引之后，族（u1，…，um，vm+1，…，vn）和（v1，…，vn）生成e的相同子空间。

证据。我们对m进行归纳，当m=0时，这个家族（u1，…，um）是空的，这个命题的意义很小。对于诱导步骤，我们有一个线性独立的家族（U1，…，Um，Um+1）。考虑线性独立族（u1，…，um）。根据诱导假设，m≤n，向量vj的m替换为（u1，…，um），这样在重命名了vs的一些指数后，家族（u1，…，um，vm+1，…，vn）和（v1，…，vn）生成了e的相同子空间。向量um+1也可以表示为一个线性组合。（v1，…，vn）和（u1，…，um，vm+1，…，vn）的组合产生了相同的子空间，因此，可以将um+1表示为（u1，…，um，vm+1，…，vn）的线性组合，例如

.

我们认为，对于一些m+1≤j≤n的j，λj=06，这意味着m+1≤n。否则，我们将

，

用户界面的一种非平凡线性依赖关系，这是不可能的，因为（u1，…，um+1）是线性无关的。

因此m+1≤n，在必要时对指数进行重命名后，我们可以假定λm+1=06，因此我们得到

.

观察族（u1，…，um，vm+1，…，vn）和（u1，…，um+1，vm+2，…，vn）生成相同的子空间，因为um+1是（u1，…，um，vm+1，…，vn）的线性组合，vm+1是（u1，…，um+1，vm+2，…，vn）的线性组合。由于（u1，…，um，vm+1，…，vn）和（v1，…，vn）生成了相同的子空间，我们得出（u1，…，um+1，vm+2，…，vn）和（v1，…，vn）生成了相同的子空间，从而得出了诱导假设的结论。

下面是一个例子，说明了替代引理。考虑序列（u1、u2、u3）和（v1、v2、v3、v4、v5），其中（u1、u2、u3）是线性独立的族，uis用vjs表示如下：

U1=v4+v5 u2=v3+v4−v5 u3=v1+v2+v3。

从我们得到的第一个方程

v4=u1−v5，

通过代入第二个方程

u2=v3+v4−v5=v3+u1−v5−v5=u1+v3−2v5。

从上面的方程我们得到

v3=−u1+u2+2v5，

如此

u3=v1+v2+v3=v1+v2−u1+u2+2v5。

最后，我们得到

v1=u1−u2+u3−v2−2v5

所以我们有

v1=u1−u2+u3−v2−2v5 v3=−u1+u2+2v5 v4=u1−v5，

这表明（u1、u2、u3、v2、v5）与（v1、v2、v3、v4、v5）跨越相同的子空间。向量（v1、v3、v4）已被（u1、u2、u3）替换，剩余的向量为（v2、v5）。我们可以重命名它们（v4、v5）。

为了完整性起见，这里是替换引理（及其证明）的更正式的声明。

提案3.8.（替换引理，第2版）给定一个向量空间e，让（ui）i∈i是e中的任意有限线性独立族，其中i=m，让（vj）j∈j是任意有限族，这样每个ui都是（vj）j∈j的线性组合，其中j=n。那么，存在一个集合l和一个注入ρ：l→j。（一个重新标号函数）使得l i，l=n−m，并且族（ui）i i（vρ（l））l l和（v j）j j生成e的相同子空间，特别是m≤n。

证据。我们从i=m开始归纳，当m=0时，族（ui）i∈i是空的，并且命题与l=j（ρ是同一性）具有平凡的关系。假设i=m+1。考虑线性独立族（ui）i∈（i−p），其中p是i的任何成员。根据归纳假设，存在一个集合l和一个注入ρ：l→j，这样l（i−p）=∅，l=n−m，并且族（ui）i∈（i−p）（vρ（l））l∈l和（vj）j∈j生成相同的e的子空间。如果p∈l，我们可以用（l−p）p0替换l，其中p0不属于i l，用与l−p上的ρ一致的注射剂ρ0替换ρ，从而使ρ0（p0）=ρ（p）。因此，我们可以一直假设l i=∅。由于up是（v j）j∈j与族（ui）i∈（i−p）（vρ（l））l∈l与（vj）j∈j的线性组合生成e的相同子空间，up是（ui）i∈（i−p）（vρ（l））l∈l的线性组合。

上=xλiui+xλlvρ（l）。（1）

I∈（I−P）L∈L

如果所有l∈l的λl=0，我们有

，

与（ui）i∈i是线性无关这一事实相矛盾。因此，对于一些l∈l，λl=06，假设l=q。由于λq=06，我们得到

（2）

我们声称家族（ui）i \123; \\\ l，by（1），vρ（q）是（ui）i∈i（vρ（l））l∈（l−q），by（2）的线性组合。因此，族（ui）i（vρ（l））l（l 123; \\ \123; \125;，（v j）j ；L−Q=N−（M+1）

其思想是，矢量vj的m可以用线性无关的ui替换，这样就可以生成相同的子空间。函数ρ：l→j的目的是选取j的n−m元素j1，…，jn−m，并将它们重新标记为l1，…，ln−m，这样这些新索引就不会与i中的索引冲突；这样，将“生存”的向量vj1，…，vjn−m重新标记为vl1，…，vln-m，其他m向量vj与j∈j−j1，…，jn−m替换为ui。这个新家族的索引集是I L。

事实上，我们可以证明，当向量空间有限生成时，命题3.8隐含定理3.5。把定理3.5和3.8放在一起，我们得到以下基本定理。

定理3.9.设e为有限生成的向量空间。任何族（ui）i∈i生成e都包含一个作为e基础的子族（uj）j∈j，任何线性独立族（ui）i∈i都可以推广到一个作为e基础的族（uj）j∈j（带有i j）。此外，对于e的每两个基（ui）i∈i和（vj）j∈j，对于某个固定整数n≥0，我们得到了i=j=n。

证据。第一部分紧接着应用定理3.5，其中l=∅和s=

（ui）i∈i.第二部分，考虑族S0=（ui）i∈i（vh）h∈h，其中（vh）h∈h是生成e的任意有限生成族，且具有i h∅。然后，将定理3.5应用于l=（ui）i∈i和S0。对于最后一个陈述，假设（ui）i∈i和（vj）j∈j是e的基，因为（ui）i∈i是线性独立的，（vj）j∈j跨越e，命题3.8意味着i≤j。对称参数产生j≤i。

注：定理3.9也适用于非有限生成的向量空间。如下所示。设（ui）i∈i为e的基，设（vj）j∈j为e的生成族，并假定i为无穷大。对于每个j∈j，设lj i为有限集

.

设l=sj∈j lj。根据定义l i，由于（ui）i∈i是e的基础，我们必须有i=l，否则（ui）i∈l将是e的另一个基础，这将与以下事实相矛盾：

（ui）il∈ij是有限的，是线性独立的。而且，我是有限的。但是，since j必须是无穷大的，否则，因为ei=sj∈j lj与j无穷大和lj有限，根据集论的标准结果，i≤j。如果（vj）j∈j也是一个基础，通过对称论证，我们得到了e的任意两个基（ui）i∈i和（vj）j∈j。

定义3.6.当一个向量空间e不是有限生成时，我们说e是无限维的。有限生成向量空间e的维数是其所有基的公共维数n，用dim（e）表示。

显然，如果把场k本身看作一个向量空间，那么a∈k和a=06的每个族（a）都是基。因此，dim（k）=1。注意dim（0）=0。

定义3.7.如果e是尺寸n≥1的向量空间，对于e的任何子空间u，如果dim（u）=1，则u称为直线；如果dim（u）=2，则u称为平面；如果dim（u）=n-1，则u称为超平面。如果dim（u）=k，那么u有时被称为k平面。

设（ui）i∈i为向量空间e的基，对于任意向量v∈e，由于族（ui）i∈i生成e，在k中有一个标量的族（λi）i∈i，这样

V=xλiui。

我爱我

一个非常重要的事实是家族（λi）i∈i是唯一的。

命题3.10.假设v=给定向量空间pi∈iλiui。那么，家族，让（ui）（iλ∈ii）是一个向量家族，i是一个标量，所以ev。let=pivi∈λieu，i

蛱蛱

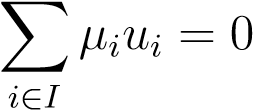
是唯一的iff（ui）i∈i是线性独立的。

证据。首先，假设（pμui）i i，那么我们有∈i是线性无关的。如果（μi）i∈i是另一个标量族

在k中，使得v=i∈i

，

由于（ui）i∈i是线性无关的，我们必须使所有i∈i的λi−μi=0，也就是说，所有i∈i的λi=μi。相反的是矛盾。如果（ui）i∈i是线性相关的，则会有一个scalars的族（祆i∈i，而不是全部为空，因此



对于一些j∈i，μj=06，但是，

V=xλi ui+0=xλiui+xμiui=x（λi+μi）ui，

i∈i i∈i i∈i i∈i

自λiu礹i.j=06以来，对于λj=6λj+p礹j，与（λi）i∈i是唯一族的假设相矛盾，使得v=i∈i

定义3.8.如果（i）i i i是向量空间E的一个基，对于任何向量v\* e，如果（Xi）i i是R中的唯一标量族，则

V=xxiui，

我爱我

每个XI被称为V的索引I的分量（或坐标）相对于基（UI）i i。

给定一个场k和任意（非空）集i，我们可以形成一个向量空间k（i），在某种意义上，它是维i\_的标准向量空间。

3.5。矩阵

定义3.9.给定一个域k和任意（非空）集i，让k（i）是由所有族（λi）i∈i组成的笛卡尔积ki的子集，其中k有有限的标量支持，我们定义了加乘和乘的标量如下：

（λi）i∈i+（μi）i∈i=（λi+μi）i∈i，

而λ·（μi）i∈i=（λμi）i∈i。

立即证明标量的加法和乘法定义良好。因此，k（i）是一个向量空间。此外，由于具有有限支撑的族是向量空间化的基础，因此向量eki的族（ei）i∈i定义为（（i）。当i=1，…，nei）j=0时，我们表示j 6=ikand（（i）byei）ki=1n。式中，是函数：i→k（i），这样，（i）=ei对于每个i∈i，显然是一个注射剂。

当我是有限集时，k（i）=ki，但当我是无限集时，这是错误的。实际上，dim（k（i））=i，但dim（ki）在无穷大的情况下更大。

许多有趣的数学结构是向量空间。一个非常重要的例子是将在下一节中定义的两个向量空间之间的一组线性映射。下面是一个例子，它将为我们准备线性映射的向量空间。

例3.5。设x为任意非空集，设e为向量空间。所有函数f:x→e的集合可以构成一个向量空间，如下所示：给定任意两个函数f:x→e和g:x→e，让（f+g）：x→e定义为：

（f+g）（x）=f（x）+g（x）

对于所有x∈x，并且对于每个λ∈k，让λf:x→e定义为：

（λf）（x）=λf（x）

对于所有的x∈x，向量空间的公理很容易被验证。

## 3.5矩阵

在第2.1节中，我们非正式地介绍了矩阵的概念。在本节中，我们将精确地定义矩阵，并介绍一些关于矩阵的操作。结果表明，矩阵形成了一个向量空间，该空间配备了一个具有关联性但非交换性的乘法运算。我们将在第4.1节中解释矩阵如何用于表示下一节中定义的线性映射。

|  |  |
| --- | --- |
| 网络错误 | |
| 网络错误 |  |
| 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 |

定义3.10.如果k=r或k=c，k上的m×n矩阵是一个族（a i j），1 i m，1 j n

在m=1的特殊情况下，我们有一个行向量，用表示

（A11···A1n）

在n=1的特殊情况下，我们有一个列向量，用

.

在最后两种情况下，我们通常省略常量索引1（行的第一个索引，列的第二个索引）。所有m×n矩阵的集合用mm、n（k）或mm、n表示。n×n矩阵称为n维数的平方矩阵。n维数的所有平方矩阵的集合用mn（k）或mn表示。

注：根据定义，矩阵a=（a i j）1≤i≤m，1≤j≤n是一个族，即从1,2，…，m×1,2，…，n到k的函数。因此，没有理由对指数进行排序。因此，通过对行或列采用不同的顺序，矩阵A可以用许多不同的方式表示为一个数组。然而，习惯上（并且通常方便）假定集合1,2，…，m和1,2，…，n的自然顺序，并根据行和列的顺序将表示为数组。

我们在矩阵上定义了一些操作，如下所示。

定义3.11.给定两个m×n矩阵a=（aij）和b=（bij），我们定义它们的和

A+B作为矩阵c=（cij），因此cij=aij+bij；也就是说，

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误 |  |  |  |
|  |  |  |  |  |  | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 |

3.5。矩阵

对于任何矩阵a=（aij），我们让−a作为矩阵（−aij）。给定一个标量λ∈k，我们将矩阵λa定义为矩阵c=（cij），这样cij=λaij；即

.

给定m×n矩阵a=（aik）和n×p矩阵b=（bkj），我们将它们的乘积ab定义为m×p矩阵c=（cij），以便

，

对于1≤i≤m和1≤j≤p，产品ab=c如下所示

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 |

注意，通过矩阵a和b相乘得到的矩阵ab的索引i和j的条目可以用a的第i行对应的行矩阵与b的j列对应的列矩阵的乘积来标识：

.

定义3.12.在维度n中，对角线上包含1，其他地方包含0的平方矩阵称为单位矩阵。它由表示

|  |  |  |  |
| --- | --- | --- | --- |
| 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误 |

定义3.13.给定m×n矩阵a=（a i j），其转置a>=（a>j i）是n×m矩阵，因此a>j i=aij，对于所有i，1≤i≤m，以及所有j，1≤j≤n。

矩阵A的转置有时用at表示，甚至用ta表示。请注意，矩阵A的转置A>具有以下特性：A>的第j行是A的第j列。换句话说，转置交换矩阵的行和列。

下面的观察将在稍后讨论SVD时有用。给定任意m×n矩阵a和任意n×p矩阵b，如果我们用a1，…，an表示a的列，用b1，…，bn表示b的行，那么我们得到

ab=a1b1+·····+anbn。

对于每一个尺寸为n的正方形矩阵a，立即验证ain=ina=a。

定义3.14.对于尺寸n的任何正方形矩阵a，如果存在a b=ba=in的矩阵b，则它是唯一的，称为a的逆矩阵。矩阵b也由a-1表示。可逆矩阵也称为非奇异矩阵，不可逆矩阵称为奇异矩阵。

利用命题3.16和矩阵表示线性映射的事实，可以证明，如果一个正方形矩阵A有一个左逆矩阵，即矩阵B，使得b a=i，或右逆矩阵，即矩阵C，使得a c=i，那么a实际上是可逆的；因此b=a−1和c=a−1。这些事实也来源于5.14号提案。

立即证明m×n矩阵的集合mm，n（k）是矩阵加上和矩阵乘上标量的向量空间。考虑m×n-矩阵Ei，j=（Ehk），定义为Eij=1，Ehk=0，如果h=6 i或k=6 j。很明显，每个矩阵a=（Aij）∈mm，n（k）都可以用一种独特的方式写为

.

因此，族（ei，j）1≤i≤m，1≤j≤n是向量空间mm，n（k）的基础，其尺寸为mn。

注：当k是一个（交换）环而不是场时，定义3.10和3.11也是完全合理的。在这个更一般的设置中，向量空间的框架太窄，但是我们可以考虑交换环a上满足定义3.1的所有公理的结构。这种结构称为模块。模理论比向量空间更为复杂。例如，模块并不总是有基础的，而向量空间的其他属性对于模块来说通常是失败的。当一个模块有基础时，它被称为自由模块。例如，当a是交换环时，结构an是一个模，使得（ei）i=1的矢量ei和（ei）j=0（j=6i）构成a的基。向量空间的许多性质仍然适用于。因此，AN是一个自由模块。例如，当a是交换环时，m m，n（a）是基（ei，j）1≤i≤m，1≤j≤n的自由模。交换环上的多项式也形成无限维的自由模。

虽然有些计算有点繁琐，但3.11号提案中列出的性质很容易得到验证。命题4.1给出了一个更具概念性的证明。

提案3.11.（1）给定任意矩阵a∈mm，n（k），b∈mn，p（k），c∈mp，q（k），我们得到

（ab）c=a（bc）；

也就是说，矩阵乘法是关联的。

（2）给定任意矩阵a，b∈mm，n（k）和c，d∈mn，p（k），对于所有的λ∈k，我们得到

（A+B）C=AC+BC

A（C+D）=AC+AD

（λa）c=λ（ac）

A（λc）=λ（ac）

使矩阵乘法·：mm，n（k）×mn，p（k）→mm，p（k）为双线性。

命题3.11的性质以及所有平方n×n矩阵的a in=in a=a的事实表明，mn（k）是一个单位在（事实上，是一个结合代数）中的环。这是一个具有零除数的非交换环，如下面的示例所示。例3.6.例如，假设a，b是2×2矩阵

然后

，

和

因此a b=6b a，ab=0，即使a，b=0.6

## 3.6线性图

既然我们了解了向量空间以及如何生成它们，我们希望能够将一个向量空间e转换为另一个向量空间f。保持向量空间结构的两个向量空间之间的函数称为向量空间的同态，或线性映射。线性映射使函数的线性概念形式化。在本节的其余部分中，我们假设所有向量空间都在给定的字段k（比如r）上。

定义3.15.给定两个向量空间e和f，e和f之间的线性映射是满足以下两个条件的函数f:e→f：

f（x+y）=f（x）+f（y）表示所有x，y∈e；f（λx）=λf（x）表示所有λ∈k，x∈e。

在第一个恒等式中设置x=y=0，我们得到f（0）=0。线性映射的基本性质是将线性组合转换为线性组合。给定e中向量的族（ui）i∈i，给定k中标量的任何族（λi）i∈i，我们得到

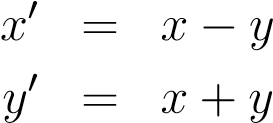
f（xλiui）=xλif（ui）。

I∈I I∈I

利用定义3.15的性质，通过对族（λiui）i∈i的支持大小的归纳，证明了上述恒等式。

例3.7.

1. 图f:r2→r2定义如下：



是一个线性映射。读者应通过以下方式检查旋转的构成

π/4，放大率√2。

1. 对于任何向量空间e，标识映射id:e→e由

id（u）=u表示所有u∈e

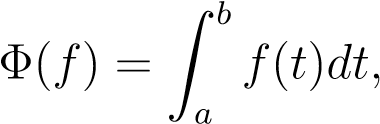
是一个线性映射。当我们想要更精确的时候，我们写的是IDE而不是ID。

1. 地图D:R[X]→R[X]定义如下：

d（f（x））=f0（x），

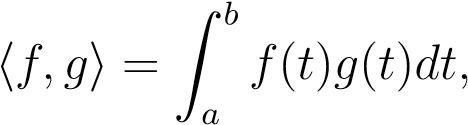
其中，f0（x）是多项式f（x）的导数，是线性映射。

1. 地图Φ：c（[a，b]）→r由



其中c（[a，b]）是在区间[a，b]上定义的一组连续函数，是一个线性映射。

1. 函数h−、−i:c（[a，b]）×c（[a，b]）→r由



在每个变量f，g中都是线性的。它还满足hf，gi=hg，fi和hf的性质，fi=0 iff=0。它是一个内部产品的例子。

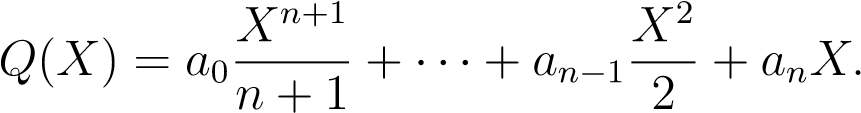
定义3.16.对于线性映射f:e→f，我们将其图像（或范围）imf=f（e）定义为集合

imf=y∈f（x∈e）（y=f（x）），

其内核（或空空间）kerf=f−1（0），作为集合

切口=x∈e f（x）=0。

例3.7（3）中的导数映射d:r[x]→r[x]的核是常数多项式，因此kerd=r。如果我们考虑二阶导数d d:r[x]→r[x]，则d d的核由度≤1的所有多项式组成。d:r[x]→r[x]的图像实际上是r[x]本身，因为每个多项式p（x）=a0xn+······+an−1x+an的次数n是n+1的多项式q（x）的导数，由



另一方面，如果我们考虑d对次数≤n的多项式的向量空间r[x]n的限制，则d的核仍然是r，但d的图像是r[x]n-1，次数≤n-1的多项式的向量空间。

提案3.12。给定线性映射f:e→f，集imf是f的子空间，集kerf是e的子空间。线性映射f:e→f是内射iff kerf=（0）（其中（0）是平凡子空间0）。

证据。对于任何x，y∈imf，有一些u，v∈e，这样x=f（u）和y=f（v），并且对于所有的λ，μ∈k，我们有

F（λu+μv）=λf（u）+μf（v）=λx+μy，

因此，λx+μy∈imf，表明imf是f的一个子空间，给定任意x，y∈kerf，我们得到f（x）=0和f（y）=0，因此，

F（λx+μy）=λf（x）+μf（y）=0，

也就是说，λx+μy∈切口，表明切口是e的一个子空间。

首先，假设切口=（0）。我们需要证明f（x）=f（y）意味着x=y。然而，如果f（x）=f（y），那么f（x）−f（y）=0，通过f的线性，我们得到f（x−y）=0。因为切口=（0），我们必须有x-y=0，也就是x=y，所以f是内射的。相反，假设f是内射的。如果x∈切口，即f（x）=0，因为f（0）=0，我们得到f（x）=f（0），通过注入率，x=0，这证明切口=（0）。因此，f是注射剂iff kerf=（0）。

由于命题3.12，线性映射f的图像imf是f的子空间，我们可以将f的秩Rk（f）定义为imf的维数。

定义3.17.对于线性映射f:e→f，f的秩rk（f）是f的图像imf的维数。

向量空间中的基的一个基本性质是，它们允许将线性映射定义为唯一的同态扩展，如下面的命题所示。

提案3.13.给定任意两个向量空间e和f，给定e的任何基（ui）i∈i，给定f中的任何其他向量族（vi）i∈i，存在一个唯一的线性映射f:e→f，使得f（ui）=vi对于所有i∈i。此外，f是内射iff（vi）i∈i是线性独立的，f是外射iff（vi）i∈i生成f。

证据。如果存在这样一个线性映射f:e→f，由于（ui）i∈i是e的基础，因此每个向量x∈e都可以唯一地作为线性组合写入。

x=xxiui，

我爱我

根据线性度，我们必须

.

定义函数f:e→f，方法是

f（x）=xxivi

我爱我

每x=pi∈i xiui。很容易验证f确实是线性的，根据前面的推理它是唯一的，显然，f（ui）=vi。

现在，假设f是内射的。设（λi）i∈i为标量的任何族，并假定

X

λivi=0.

我爱我

因为每个i∈i的vi=f（ui），我们有

f（xλiui）=xλ，如果（ui）=xλivi=0。

i∈i i∈i i∈i

既然f是注射剂iff kerf=（0），我们有

，

由于（ui）i∈i是一个基，我们得到所有i∈i的λi=0，这表明（vi）i∈i是线性无关的。相反，假设（e中，每个向量x∈e是一个线性组合vi）i∈i是线性独立的。因为（x=pi∈iλi ui of（ui）i∈i.如果ui）i∈i是一个基础

，

然后

xxx

λivi=λif（ui）=f（λiui）=0，

i∈i i∈i i∈i

所有i∈i的λi=0，因为（vi）i∈i是线性无关的，这意味着x=0。因此，切口=（0），这意味着f是内射的。其中f是主观性的部分作为一个简单的练习。

在3.13命题的第二部分，一个内射线性映射f:e→f向一个线性独立族（f（ui））发送一个基（ui）i∈i，当f为双射时，它也是一个基。当e和f具有相同的有限维n时，（ui）i∈i是e的基，f:e→f是内射的，（f（ui））i∈i是f的基（由命题3.6）。

现在我们可以证明定义3.9的向量空间k（i）具有一个普适性，等于说k（i）是i自由生成的向量空间。回想一下，：i→k（i），这样，对于每个i∈i，（i）=ei是从i到k（i）的注入。

提案3.14.对于任意向量空间f和任意函数f:i→f，

有一个独特的线性图f:k（i）→f，这样

F=F\_，

如下图所示：

我是CCCC/CK！（FI）

f

f

证据。如果存在这样一个线性映射f:k（i）→f，因为f=f\_

f（i）=f（（i））=f（ei），

向量信息每一个if∈，并由命题3.13，有一个唯一的线性mapi。然而，家庭（ei）i∈i是k（i）的基础，（ff：（ki））（ii）∈i→是一个家庭关闭这样

f（ei）=f（i）对于每一个i∈i，证明了线性映射f的存在唯一性

这样f=f\_。

下面的简单命题也很有用。

提案3.15。给定任意两个向量空间e和f，其中f非平凡，给定e中向量的任意族（ui）i∈i，则以下性质成立：

1. 族（ui）i∈i在f中为每个向量族（vi）i∈i生成e iff，其中至多有一个线性映射f:e→f，这样f（ui）=vi代表所有i∈i。
2. 族（ui）i∈i是线性独立的i f f，对于每个向量族（vi）i∈i，在f中有一些线性映射f:e→f，这样f（ui）=vi代表所有i∈i。

证据。（1）如果有任何线性映射f:e→f，使得f（ui）=vi对于所有i∈i，因为（ui）i∈i生成e，每个向量x∈e都可以写成某种线性组合。

x=xxiui，

我爱我

根据线性度，我们必须

f（x）=xxif（ui）=xxivi.

I∈I I∈I

这表明f是唯一的，如果它存在的话。相反，假设（ui）i∈i不生成e，因为f是非平凡的，所以有一些向量y∈f，使得y=06。由于（ui）i∈i不生成e，在（ui）i∈i生成的子空间中有一个不存在的向量w∈e，根据定理3.9，在（ui）i∈i生成相同子空间时有一个线性独立的子族（ui）i∈i0。由于假设w∈e不在（ui）i∈i0，由引理3.4和定理3.9生成的子空间中，又有e的一个基（e j）j∈i0 j，使得ei=ui，对于所有i∈i0和w=e j0，对于一些j0∈j，让（vi）i∈i成为f中的族，这样，vi=0代表所有i∈i，defin当f:e→f为0的常线性映射时，我们有一个线性映射，使f（ui）=0代表所有i∈i。根据命题3.13，有一个唯一的线性映射g:e→f，这样g（w）=y，g（e j）=0代表所有j∈（i0 j）−j0。根据e的基（e j）j i0 j的定义，对于所有i i，我们都有，g（ui）=0，并且由于f 6=g，这与这样一个映射最多存在一个的事实相矛盾。

（2）如果族（ui）i∈i是线性独立的，那么根据定理3.9，（ui）i∈i可以推广到e的基，其结论由命题3.13得出。相反，假设（ui）i∈i是线性相关的。那么，有一个标量（并非全部为零）的族（λi）i∈i，这样

X

λiui=0.

我爱我

假设任意非零向量y∈f，对于每一个i∈i，都有一个线性映射fi:e→f，这样fi（ui）=y，fi（uj）=0，对于j∈i−i。那么，我们就可以

0=fi（xλiui）=xλifi（ui）=λiy，

I∈I I∈I

由于y=06，这意味着λi=0，对于每一个i∈i，因此，（ui）i∈i是线性无关的。

在给定向量空间e、f和g以及线性映射f:e→f和g:f→g的情况下，很容易证明f和g的组合g f:e→g是线性映射。

定义3.18。线性映射f:e→f是同构的，如果存在线性映射g:f→e，那么

G F=IDE和F G=IDF。（）

定义3.18中的地图G是唯一的。这是因为如果g和h都满足g\_f=ide、f\_g=idf、h\_f=ide和f\_h=idf，那么

g=g idf=g（f h）=（g f）h=ide h=h。

上面满足（）的图G称为f的倒数，它也用f-1表示。

命题3.13意味着如果e和f是两个向量空间，（ui）i∈i是e的基，f:e→f是同构的线性映射，那么族（f（ui））i∈i是f的基。

我们可以验证，如果f:e→f是一个双射线性映射，那么它的逆f−1:f→e也是一个线性映射，因此f是同构的。

命题3.13的另一个有用推论是：

提案3.16。设e为有限维n≥1的向量空间，设f:e→e为任意线性映射。以下属性保留：

1. 如果f有一个左逆g，也就是说，如果g是一个线性映射，这样g f=id，那么f是同构的，f−1=g。
2. 如果f有一个右逆h，也就是说，如果h是一个线性映射，这样f h=id，那么f是同构，f−1=h。

证据。（1）方程g f=id表示f是内射的；这是关于函数的标准结果（如果f（x）=f（y），那么g（f（x））=g（f（y）），这意味着自g f=id以来x=y）。假设（u1，…，un）是e的任何基。根据命题3.13，因为f是内射的，（f（u1），…，f（un））是线性独立的，并且由于e有维n，所以它是e的基（如果（f（u1），…，f（un））不跨越e，那么它可以严格地扩展到维的基大于n，这与OREM 3.9）。然后，F是双射的，根据先前的观察，它的逆是一个线性图。我们也有

G=G ID=G（F F−1）=（G F）F−1=ID F−1=F−1。

（2）方程f h=id表示f是可射的；这是关于函数的标准结果（对于任意y∈e，我们有f（h（y））=y）。假设（u1，…，un）是e的任何基础。根据命题3.13，因为f是主观性的，（f（u1），…，f（un））跨越e，并且因为e有维n，所以它是e的基础（如果（f（u1），…，f（un））不是线性独立的，那么因为它跨越e，它包含的维的基础严格小于n，矛盾ng定理3.9）。然后，F是双射的，根据先前的观察，它的逆是一个线性图。我们也有

H=ID H=（F−1 F）H=F−1（F H）=F−1 ID=F−1。

这就完成了证明。

定义3.19.两个向量空间e和f之间的所有线性映射集都用hom（e，f）或l（e；f）表示（符号l（e；f）通常保留给连续线性映射集，其中e和f是赋范向量空间）。当我们想要更精确地定义向量空间e和f的K域时，我们写homk（e，f）。

集合hom（e，f）是在节末定义的操作下的向量空间。

2.1，即

（f+g）（x）=f（x）+g（x）

对于所有x∈e，和

（λf）（x）=λf（x）

对于所有x e，值得仔细检查的一点是，λf确实是一个线性映射，它使用了k域中的交换性。事实上，我们有

（λf）（μx）=λf（μx）=λμf（x）=λf（x）=μ（λf）（x）。

当e和f有有限维时，向量空间hom（e，f）也有有限维，稍后我们将看到。

定义3.20。当e=f时，线性映射f:e→e也被称为自同态。空间hom（e，e）也用end（e）表示。

同样重要的是要注意的是，组成赋予了HOM（e，e）一个环结构。实际上，组合是一个操作：hom（e，e）×hom（e，e）→hom（e，e），它是关联的，具有标识ide，并且分布性属性保持：

（g1+g2）f=g1 f+g2 f；g（f1+f2）=g f1+g f2。

环hom（e，e）是非交换环的一个例子。

很容易看出，f:e→e的双射线性映射集是一个组合下的群。

定义3.21。双射线性映射f:e→e也称为自同构。e的自同构群称为一般线性群（e），用gl（e）或aut（e）表示，当e=r n时，用gl（n，r）表示，甚至用gl（n）表示。

虽然在这本书中，我们不会有很多场合使用商空间，但它们是代数的基础。下一节可以省略，直到需要为止。

3.7。商空间

## 3.7商空间

设e为向量空间，设m为e的任意子空间，子空间m引出一个关系式

\_m on e，定义如下：对于所有u，v∈e，u m v iff u−v∈m。

我们有以下简单的建议。

提案3.17。对于任意向量空间e和e的任意子空间m，关系m是与以下两个同余性质的等价关系：

1. 如果U1 M v1和U2 M v2，则U1+U2 M v1+v2，以及
2. 如果u\_m v，则λu\_mλv.

证据。很明显，m是等价关系。注意，u1 m v1和u2 m v2等于u1−v1=w1和u2−v2=w2，其中w1、w2∈m，因此，

（u1+u2）−（v1+v2）=w1+w2，

而w1+w2∈m，因为m是e的子空间，所以我们得到u1+u2 m v1+v2。如果u−v=w，用w∈m，则λu−λv=λw，

而λw∈m，因为m是e的子空间，因此λu mλv。

命题3.17表明，我们可以在等价关系m的等价类的集合e/m上定义一个标量的加法和乘法。

定义3.22.给定任意向量空间e和e的任意子空间m，在等价关系m的等价类集合e/m上，我们定义了一个标量的加乘运算，如下所示：对于任意两个等价类[u]，[v]∈e/m，我们得到

[U]+[V]=[U+V]，λ[U]=[λU]。

在命题3.17中，上述运算不依赖于等价类[u]、[v]∈e/m中代表的具体选择，直接验证e/m是一个向量空间。函数π：e→e/f，定义为π（u）=[u]对于每个u∈e，是一个称为e对e/f的自然投影的投影的投影线性映射，向量空间e/m被子空间m称为e的商空间。

对于任何线性映射f:e→f，我们知道kerf是e的一个子空间，并立即证明imf与商空间e/kerf同构。

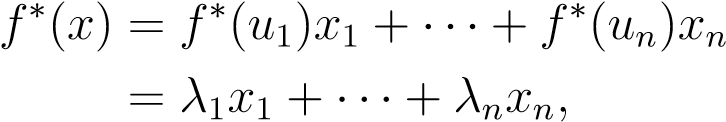
## 3.8线性形式和对偶空间

我们已经观察到场k本身（k=r或k=c）是一个向量空间（在其自身之上）。从E到K场的线性映射的向量空间hom（e，k），即线性形式，起着特殊的作用。在这一部分中，我们只定义线性形式，并证明每个有限维向量空间都有一个对偶基。第10章给出了对偶空间和对偶性的更高级的表示。

定义3.23。对于向量空间e，从e到场k的线性映射的向量空间hom（e，k）称为e的对偶空间（或对偶空间）。空间hom（e，k）也用e表示，e中的线性映射称为线性形式，或covector。空间e的双空间e称为e的双空间。

记法上，线性形式f:e→k也用星号表示，如u、x等。

如果e是有限维n的向量空间，并且（u1，…，un）是e的基础，对于任何线性形式f∈e，对于每个x=x1u1+······+xnun∈e，通过线性我们得到

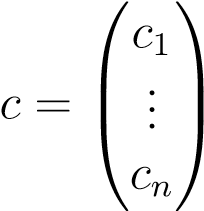


当λi=f（ui）k为每i，1≤i≤n时，因此，对于基（u1，…，un），线性形式f由行向量表示。

我们有

，

x坐标的线性组合，我们可以把线性形式f看作一个线性方程。如果我们决定使用系数的列向量



而不是行向量，则线性形式f定义为

f（x）=c>x。

上面的符号通常用于机器学习。

3.8。线性形式与对偶空间

例3.8。给定任意可微函数f:rn→r，根据定义，对于任意x∈rn，f在x处的总导数dfx是线性形式dfx:rn→r，因此对于所有u=（u1，…，un）∈rn，

.

例3.9.设c（[0,1]）为连续函数f[0,1]→r的向量空间，映射i:c（[0,1]）→r由

对于任意f∈c（[0,1]）

是线性形式（积分）。

例3.10。考虑实n×n矩阵的向量空间mn（r）。设tr:mn（r）→r为tr（a）=a11+a22+·····+ann给出的函数，

称为A的迹线。它是一种线性形式。设为：mn（r）→r为

，

其中a=（aij）。立即证明S是线性形式。

给定一个向量空间e和任意基（u i）i∈i for e，我们可以将一个线性形式u i∈e与每个ui关联，并且u i具有一些显著的性质。

定义3.24.给定一个向量空间e和任意基（ui）i∈i for e，通过命题3.13，对于每一个i∈i，都有一个唯一的线性形式，这样

，

对于每个j∈i，线性形式称为索引i w.r.t的坐标形式。基础

（ui）i∈i。

注：对于指数集i，作者通常定义所谓的“克罗内克符号”δij，以便

J J，

对于所有i，j∈i，那么，u i（uj）=δij。

术语坐标形式的原因如下：如果e有有限维，如果（u1，…，un）是e的基础，对于任何向量。

V=λ1u1+····+λnun，

我们有

u i（v）=u i（λ1u1+····+λnun）

=λ1u i（u1）+····+λiu i（ui）+···+λnu i（un）=λi，

因为u i（uj）=δij。因此，u i是一个线性函数，它返回在基上表示的向量的第i个坐标（u1，…，un）。

下面的定理表明，在有限维中，向量空间e的每一个基（u1，…，un）都产生了对偶空间e的基（），称为对偶基。

定理3.18。（双基的存在）设e为维数n的向量空间，其性质如下：对于e的每一个基（u1，…，un），坐标形式族是e的基（称为（u1，…，un）的双基）。证据。（a）如果v∈e是任何线性形式，考虑线性形式

.

因为u i（uj）=δij，

f（ui）=（v（u1）u 1+·····+v（u n）u（ui）

=v（u1）u 1（u i）+····+v（ui）u i（ui）+····+v（u n）u n（ui）

=v（ui）

因此F和V\_在基础上（U1，…，Un）达成一致，这意味着

.

因此，（）跨度e。我们声称这些弯曲器是线性独立的。如果不是，我们有一个非平凡的线性依赖

，

如果我们将上述线性形式应用到每个用户界面，使用一个家族计算，我们得到

0=λiu i（ui）=λi，

证明这确实是线性无关的。因此，（）是

E.

特别地，定理3.18给出了一个有限维向量空间，它的对偶E具有相同的维数。

3.9。总结

## 3.9总结

本章的主要概念和结果如下：

* 组、环和字段。
* 向量空间的概念。
* 病媒家族。
* 矢量的线性组合；一个族的线性依赖性和线性独立性
* 线性子空间。
* 子空间跨度。（或生成）族；生成器，有限生成的子空间；基
* 每个线性独立的族都可以扩展到一个基（定理3.5）。
* 一个向量的最小生成族（命题3.6），b是一个基，iff是一个极大的线性独立族，iff是
* 替代引理（建议3.8）。
* 这是e（定理3.9）有限生成向量空间维中的任意两个基，e具有相同数量的元素；
* 超平面。
* 每一个向量在一个基上都有一个唯一的表示（以坐标表示）。
* 线性映射的概念。
* 线性映射f的图像imf（或范围）。
* 线性映射f的核切口（或空空间）。
* 线性映射f的秩Rk（f）。
* 线性映射的图像和核心是子空间。线性映射是内射的，如果它的核心是平凡空间（0）（命题3.12）。
* 关于基的线性映射的（命题唯一同态扩展属性3.13）。
* 商空间。
* 线性形式（covector）和双空间e。
* 坐标形式。
* 双基的存在（有限维）。

第四章

# 矩阵和线性映射

## 4.1用矩阵表示线性图

命题3.13表明，给定两个向量空间e和f以及e的基（uj）j∈j，

基向量（f:e→f）的每一个线性映射f下的任意向量空间之间的同构由族（uj）j∈ej唯一确定。因此，特别是考虑到尺寸j和k（jf）。如果（ufj））=j j∈kj=图像的（j）1，我们得到，…，n，a

维度n的向量空间e与向量空间kn同构。

如果我们也有f的基（vi）i∈i，那么每个向量f（uj）都可以用一种独特的方式写成

F（UJ）=Xaijvi，

我爱我

式中j∈j，对于一个标量族（aij）i∈i，因此，对于e的两个基（uj）j∈j和f的（vi）i∈i，线性映射f完全由一个可能无限的“i×j-矩阵”m（f）=（aij）i∈i，j∈j确定。

注：注意，我们故意将索引集j赋给e的基（uj）j∈j，将索引集i赋给f的基（vi）i∈i，这样，与f:e→f相关的矩阵m（f）的行被i索引，矩阵m（f）的列被j索引。显然，这导致了令人不愉快的逆转。如果我们考虑e的基（ui）i∈i和f的基（vj）j∈j，我们将得到一个j×i矩阵m（f）=（aj i）j∈j，i∈i，不管我们做什么，都会有一个反转！我们决定坚持e的基（uj）j∈j和f的基（vi）i∈i，这样我们得到一个i×j矩阵m（f），知道我们可能偶尔会遇到这个决定！

当i和j是有限的，也就是说，当i\_=m和j=n时，线性映射f由矩阵m（f）确定，其中j列中的条目是

八十三

基（v1，…，vm）上的向量f（uj），即矩阵

|  |  |  |  |
| --- | --- | --- | --- |
| 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 |

第i行和第j列的条目为aij（1≤i≤m，1≤j≤n）。

现在我们将证明，当e和f有有限维时，线性映射可以很方便地用矩阵表示，并且线性映射的组合对应于矩阵乘法。由于埃米尔·阿汀的缘故，我们将非常密切地遵循一种优雅的呈现方法。

设e和f为两个向量空间，假设e有一个有限基（u1，…，un），f有一个有限基（v1，…，vm）。回想一下，我们已经证明，每个向量x∈e都可以用一种独特的方式写成

1. =x1u1+·····+xnun，

同样地，每一个向量y∈f都可以用一种独特的方式写成

1. =y1v1+·····+ymvm。

设f:e→f为e与f之间的线性映射，那么，对于e中的每一个x=x1u1+·····+xnun，通过线性，我们得到f（x）=x1f（u1）+·····+xnf（un）。

让

，

或者更简明扼要地说，

对于每个j，1≤j≤n。这可以通过将f（uj）的系数a1j，a2j，…，amj写在基（v1，…，vm）上，作为矩阵的jth列来表示，如下所示：

F（U1）F（U2）…F（UN）

v1 a11 a12…A1N V2 A21 A22…A2N\_

……………\_\_

虚拟机AM1 AM2…AMN

然后，将每个f（uj）的右边代入f（x）的表达式，我们得到

，

它通过重新组合项来获得vi的线性组合，从而得出

.

因此，假设f（x）=y=y1v1+····+ymvm，我们得到

（1）

对于所有i，1≤i≤m。

为了使事情更具体，让我们来处理n=3和m=2的情况。在这种情况下，

F（U1）=A11v1+A21v2 F（U2）=A12v1+A22v2 F（U3）=A13v1+A23v2，

以矩阵形式表示为

，

对于任何x=x1u1+x2u2+x3u3，我们有

f（x）=f（x1u1+x2u2+x3u3）

=x1f（u1）+x2f（u2）+x3f（u3）

=x1（a11v1+a21v2）+x2（a12v1+a2v2）+x3（a13v1+a23v2）=（a11x1+a12x2+a13x3）v1+（a21x1+a2x2+a23x3）v2。

因此，由于y=y1v1+y2v2，

我们有

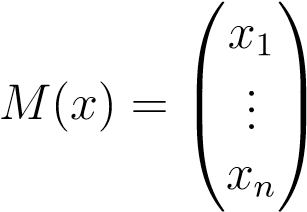
y1=a11x1+a12x2+a13x3 y2=a21x1+a2x2+a23x3。

这与矩阵方程一致。

.

现在我们用矩阵形式化线性映射的表示。

定义4.1.设e和f为两个向量空间，设（u1，…，un）为e的基，（v1，…，vm）为f的基，每个向量x∈e在基（u1，…，un）中表示为x=x1u1+·····+xnun由列矩阵表示



同样地，对于基（v1，…，vm）中表示的每个向量y∈f。

每个线性映射f:e→f由矩阵m（f）=（aij）表示，其中aij是基（v1，…，vm）上向量f（uj）的第i个分量，即

，每j，1≤j≤n。

基础（v1，…，vm）上f（uj）的系数a1j，a2j，…，amj构成

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 网络错误 |  |  |  |  |
|  | 网络错误 | 网络错误 | 网络错误 | 网络错误 |
| 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 |

与线性映射f:e→f相关的矩阵m（f）称为f相对于基（u1，…，un）和（v1，…，vm）的矩阵。当e=f且基（v1，…，vm）与e的基（u1，…，un）相同时，与f:e→e（如上）相关联的矩阵m（f）称为f相对于基（u1，…，un）的矩阵。

备注：在定义3.10之后的备注中，没有理由假定基（u1，…，un）和（v1，…，vm）中的向量是以任何特定方式排序的。然而，假设自然顺序通常很方便。如果是这样，作者有时将矩阵m（f）称为关于有序基（u1，…，un）和（v1，…，vm）的f矩阵。

现在让我们考虑如何用基来表示线性映射的组成。

设e、f和g为三个向量空间，其中e的基（u1，…，up），f的基（v1，…，vn），g的基（w1，…，wm）。设g:e→f和f:f→g为线性映射。如前所述，G:E→F由基向量UJ的图像决定，F:F→G由基向量VK的图像决定。我们想了解f g:e→g是如何由基向量uj的图像确定的。

备注：请注意，我们正在考虑线性映射g:e→f和f:f→g，而不是f:e→f和g:f→g，这将生成f g:e→g而不是g f:e→g的组合。我们可能不寻常的选择是由以下事实驱动的：如果f由矩阵m（f）=（aik）和g表示用矩阵m（g）=（bkj）表示，则f g:e→g用矩阵a和b的乘积ab表示。如果我们采用了另一种选择，其中f:e→f和g:f→g，则g f:e→g用乘积ba表示。就个人而言，当两个矩阵的积是AB而不是BA写的时候，我们发现更容易记住这两个矩阵积的j行和j列中的条目公式。显然，这是一个品味问题！我们将不得不接受我们可能非正统的选择。

因此，让

，

对于每k，1≤k≤n，并让

，

对于每个j，1≤j≤p；在矩阵形式中，我们有

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 网络错误 | 网络错误 | 网络错误 | 网络错误 |
| 网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 |
|  | 网络错误 | 网络错误 | 网络错误 | 网络错误 |
| 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误 |
| 网络错误 | |

1. =x1u1+·····+xpup，

假设g（x）=y=y1v1+····+ynvn，我们有

（2）

对于所有K，1≤K≤N，以及

1. =y1v1+·····+ynvn，

假设f（y）=z=z1w1+····+zmwm，我们有

（3）

对于所有的i，1≤i≤m。那么，如果y=g（x）和z=f（y），我们有z=f（g（x）），鉴于（2）和（3），我们有

.

因此，将CIJ定义为

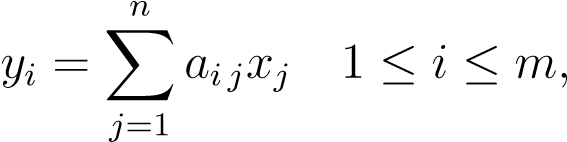
，

对于1≤i≤m和1≤j≤p，我们有

（4）

恒等式（4）表明，线性映射的合成对应于矩阵的乘积。

然后，给出用矩阵m（f）=（aij）w.r.t表示的线性映射f:e→f。用方程（1）表示基（u1，…，un）和（v1，…，vm），即



矩阵乘法的定义，方程y=f（x）对应于矩阵方程m（y）=m（f）m（x），也就是说，

.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 网络错误 |  |  |  |  |  |
| 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误 |

有时，需要将表示f的矩阵m（f）的符号中的基（u1，…，un）和（v1，…，vm）与这些基结合起来。结果这是一个混乱的企业！

我们建议采取以下行动：

定义4.2.对于e和f的基，写u=（u1，…，un）和v=（v1，…，v m），并用mu，v（f）表示f相对于u和v的矩阵。此外，对于x的坐标m（x）=（x1，…，xn）写xu，对于y的坐标m（y）=（y1，…，xn）写yv。T.基础V。然后，

y=f（x）

以矩阵形式表示为yv=mu，v（f）xu。

当u=v时，我们将mu，v（f）缩写为mu（f）。

上述表示法看似合理，但有一点缺点，即在表达式mu，v（f）xu中，输入到矩阵mu，v（f）的参数xu不出现在mu，v（f）的下标u旁边。我们可以使用符号mv，u（f），有些人会这样做。但是，我们发现有点混淆，当f从e空间映射到f空间时，v在u之前，所以我们更喜欢使用符号mu，v（f）。

请注意，其他作者（如Meyer[122]使用符号[f]u、v，其他作者（如Dummit和Foote[55]）使用符号mv（f），而不是mu、v（f）。情况越来越糟！

U

您可以找到符号mvu（f）（如lang[106]所示），或u[f]v，或其他奇怪的符号。

让我们用一个矩阵来说明具体情况下线性图的表示。设e为次数最多为4的多项式的向量空间r[x]4，设f为次数最多为3的多项式的向量空间r[x]3，且线性映射为导数映射d：即，

d（p+q）=dp+dq d（λp）=λdp，

我们选择（1，x，x2，x3，x4）作为E和（1，x，x2，x3）的基础，作为f的基础，然后通过在1（x，x，x2，x3）的基础上，分别表示i＝0，1，2，3，4的每个基向量Xi的导数DXI，得到与D相关的4×5矩阵D。我们发现

.

那么，如果p表示多项式

P=3x4−5x3+x2−7x+5，

我们有

dp=12x3−15x2+2x−7，

多项式p由矢量（5、−7,1、−5,3）表示，dp由矢量（−7,2、−15,12）表示，我们有

，

如预期的那样！D的核（空空间）由度0的多项式组成，即常数多项式。因此，dim（kerd）=1，从

dim（e）=dim（kerd）+dim（imd）

（见定理5.11），我们得到dim（imd）=4（因为dim（e）=5）。

有趣的是，让我们从向量空间r[x] 3中找出线性映射到向量空间r[x] 4，由i（i，0，1，2，3）的积分（求基元，或反导数）给出。代表R的5×4矩阵与之前相同的基是

0 0 0 0\_

10 0 0\_

S=0 1/2 0 0。

 

0 0 1/3 0\_

0 0 0 1/4

我们证实ds=i4，

，

应该的！方程d s=i4表明s是内射的，d是左逆的。但是，sd=6 I5，而不是

，

因为常数多项式（度为0的多项式）属于D的核心。

将矩阵m（f）w.r.t.的基（u1，…，un）和（v1，…，vm）关联到线性映射f:e→f的函数具有矩阵乘法对应于线性映射组合的性质。这允许我们将线性映射的属性转移到矩阵中。下面是这种技术的一个例子：

提案4.1.（1）给定任意矩阵a∈mm，n（k），b∈mn，p（k），c∈mp，q（k），我们得到

（ab）c=a（bc）；

也就是说，矩阵乘法是关联的。

（2）给定任意矩阵a，b∈mm，n（k）和c，d∈mn，p（k），对于所有的λ∈k，我们得到

（A+B）C=AC+BC

A（C+D）=AC+AD

（λa）c=λ（ac）

A（λc）=λ（ac）

使矩阵乘法·：mm，n（k）×mn，p（k）→mm，p（k）为双线性。证据。（1）每个m×n矩阵a=（aij）定义函数fa:kn→km，由

fa（x）=ax，

对于所有x∈kn。立即证明fa是线性的，用kn和km表示fa的矩阵m（fa）等于a，公式（4）证明

m（fa\_fb）=m（fa）m（fb）=ab，

所以我们得到

m（（fa fb）fc）=m（fa fb）m（fc）=（ab）c

和

m（fa（fb\_fc））=m（fa）m（fb\_fc）=a（bc）

由于函数的组合是关联的，所以我们有（fa fb）fc=fa（fb fc），这意味着

（ab）c=a（bc）。

（2）立即证实，如果f1，f2∈homk（e，f），a，b∈mm，n（k），（u1，…，un）是e的任何基，（v1，…，vm）是f的任何基，则

m（f1+f2）=m（f1）+m（f2）fa+b=fa+fb。

然后我们有了

（a+b）c=m（fa+b）m（fc）

=m（fa+b\_fc）

=m（（fa+fb）fc）

=m（（fa\_fc）+（fb\_fc））。

=m（fa\_fc）+m（fb\_fc）

=m（fa）m（fc）+m（fb）m（fc）=ac+bc。

用相似的方法证明了方程A（c+d）=ac+ad，最后两个方程很容易得到验证。我们也可以通过矩阵计算来验证所有的恒等式。

注意，命题4.1意味着方阵的向量空间mn（k）是

（非交换）单位在的环。（它甚至表明mn（k）是一个结合代数。）

下面的命题说明了hom（e，f）与m m，n之间映射f→7 m（f）的主要性质，简而言之，它是向量空间的同构。

提案4.2.给定三个向量空间e，f，g及其各自的基（u1，…，up），（v1，…，vn）和（w1，…，wm），将矩阵m（g）关联到线性映射g:e→f的映射m:hom（e，f）→mn，p满足所有x∈e，all g，h:e→f和all f:f→g的以下性质：

m（g（x））=m（g）m（x）

m（g+h）=m（g）+m（h）m（λg）=λm（g）

m（f\_g）=m（f）m（g）

4.2。基矩阵的变化

其中m（x）是与向量x相关联的列向量，m（g（x））是与g（x）相关联的列向量，如定义4.1所述。

因此，m:hom（e，f）→m n，p是向量空间的同构，当p=n且基（v1，…，vn）与基（u1，…，up）相同时，m:hom（e，e）→mn是环的同构。

证据。m（g（x））=m（g）m（x）是在陈述命题之前用同一性（1）表示的。等式m（g+h）=m（g）+m（h）和m（λg）=λm（g）是直接的，m（f\_g）=m（f）m（g）遵循（4）和矩阵乘法的定义。映射m:hom（e，f）→mn，p显然是内射的，因为每个矩阵定义了一个线性映射（见命题4.1），它也是内射的，因此是双射的。鉴于上述恒等式，它是同构的（与m:hom（e，e）→mn相似，其中命题4.1用于表示mn是一个环）。

从命题4.2来看，将有限维向量空间中的向量表示为列向量而不是行向量似乎更可取。因此，从现在开始，我们将RN的向量（或者更一般地说，kn的向量）表示为columm向量。

## 4.2基矩阵的变化

重要的是要注意，命题4.2给出的同构m:hom（e，f）→mn，p取决于碱基（u1，…，up）和（v1，…，vn）的选择，同构m:hom（e，e）→mn也同样取决于碱基（u1，…，un）的选择。因此，了解基的变化如何影响线性映射f:e→f作为矩阵的表示是很有用的。需要以下简单的建议。

提案4.3.设e为向量空间，设（u1，…，un）为e的基础。对于每个族（v1，…，vn），设p=（aij）为定义的矩阵。矩阵p是可逆的，iff（v1，…，vn）是e的基础。

证据。注意，我们有p=m（f），与唯一线性映射相关的矩阵f:e→e，这样f（ui）=vi。根据命题3.13，f是双目标iff（v1，…，vn）是e的基础。此外，很明显，中的单位矩阵是与单位id相关的矩阵：e→e w.r.t.an。Y基。如果f是同构，那么f f−1=f−1 f=id，根据命题4.2，我们得到m（f）m（f−1）=m（f−1）m（f）=in，表明p是可逆的，m（f−1）=p−1。

提案4.3提出了以下定义。

定义4.3.给定一个维数为n的向量空间e，对于e的任意两个基（u1，…，un）和（v1，…，vn），设p=（aij）为可逆矩阵，定义如下：

，

也就是关于碱基（v1，…，vn）和（u1，…，un）的单位ID的矩阵：e→e。实际上，我们在基（u1，…，un）上表示每个id（vj）=vj。

VJ的系数a1j，a2j，…，aj在基（u1，…，un）上形成了

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 网络错误 |  |  |  |  |
|  | 网络错误 | 网络错误 | 网络错误 | 网络错误 |
| 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 |

矩阵p被称为基矩阵从（u1，…，un）到（v1，…，vn）的变化。

显然，基矩阵从（v1，…，vn）到（u1，…，un）的变化是p-1。因为p=（aij）是关于基（v1，…，vn）和

（u1，…

在第4.2号提案的基础上（v1，…，vn），我们有

，

表明X（过）（U1，…，Un）的旧坐标（XI）用新坐标表示（（V1，…，VN））。

现在，我们面临着一个痛苦的任务，分配一个“好”的符号，把基u=（u1，…，un）和v=（v1，…，vn）合并到基矩阵从u变为v的符号中。因为基矩阵从u变为v是关于基v的标识映射ide的矩阵。按照这个顺序，我们可以用mv，u（id）来表示它（meyer[122]使用符号[i]v，u）。我们更喜欢使用mv，u（id）的缩写。

定义4.4.表示基矩阵从u到v的变化。

U.

注意

.

那么，如果我们写x u=（x1，…，xn），表示x相对于u的旧坐标，和）表示x相对于v的新坐标，我们得到

xu=pv，u xv，xv=pv−，1u xu。

上面可能回顾一下，但要记住，矩阵mu，v（f）接受以U为基础表示的输入，以V为基础表示的输出。因此，pv，u接受以V为基础表示的输入，以U为基础表示的输出，xu=pv，u xv符合这一观点！

4.2。基矩阵的变化

注意一些作者（如Artin[7]）定义了基矩阵从u到的变化。在这个观点下，旧的基础u用新的基础v表示，我们发现这有点不自然。而且，在实践中，似乎新的基础经常用旧的基础来表达，而不是用另一种方式来表达。

由于矩阵P＝PV，U在旧基（U1，…，Un）的基础上表达了新的基（V1，…，VN），我们观察到矢量X的坐标（XI）在基变化的相反方向上变化。由于这个原因，向量有时被认为是反变的。然而，这个表达没有意义！实际上，不依赖于特定基础的内在量的向量。有意义的是，矢量的坐标以一种反变的方式变化。

让我们考虑一些改变基础的具体例子。

例4.1.设e=f=r2，其中u1=（1,0），u2=（0,1），v1=（1,1）和v2=（-1,1）。基矩阵p从基u=（u1，u2）到基v=（v1，v2）的变化是

它的反方向是

.

-

相对于（u1，u2）的旧坐标（x1，x2）表示为相对于（v1，v2）的新坐标（）。

，

关于（v1，v2）的新坐标（）用关于（u1，u2）的旧坐标（x1，x2）表示为

.

例4.2.设e=f=r[x]3为次数最多为3的多项式集，并考虑基U=（1，x，x2，x3）和），其中

）是3阶的伯恩斯坦多项式，由

.

通过扩展伯恩斯坦多项式，我们发现基矩阵pv，u的变化由

.

我们还发现pv，u的倒数是

.

因此，基于v的多项式2x3−x+1的坐标为

.

我们的下一个例子是haar小波，它是信号处理的基本工具。

## 4.3 Haar基向量和小波的一瞥

我们首先考虑r4中的haar小波。小波在音频和视频信号处理中起着重要的作用，尤其是对于将长信号压缩成比保留足够信息小得多的信号，这样当它们被播放时，我们就看不到或听不到任何区别。

考虑四个向量w1，w2，w3，w4，由

.

请注意，这些向量是成对正交的，因此它们确实是线性无关的（我们将在后面的章节中看到这一点）。设w=w1，w2，w3，w4为haar基，设u=e1，e2，e3，e4为r4的规范基。基矩阵w=pw，u从

u到w由

，

−−

我们很容易发现w的倒数由

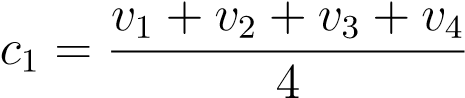
.

因此，u基上的向量v=（6,4,5,1）在haar基上变成c=（c1，c2，c3，c4）

W，带

.

给定一个信号v=（v1，v2，v3，v4），我们首先通过计算c=w−1v将v转换为其系数c=（c1，c2，c3，c4）。观察



是信号v的总平均值。系数c1对应于图像（或声音）的背景。然后，c2给出v的粗略细节，而c3给出v的第一部分的细节，c4给出v的下半部分的细节。

信号重建包括计算v=wc。好的压缩技巧是去掉C的一些系数（设置为零），获得一个压缩的信号C，并且仍然保留足够的关键信息，以便重建的信号V=wc b b b看起来几乎和原始信号V一样好。因此，步骤是：输入V−→系数c=w−1v−→压缩c−→压缩v=wc。B-B

这种压缩方案使现代视频会议成为可能。

事实证明，在不实际使用w-1的情况下，有一种更快的方法可以找到c=w-1v。这与Haar小波的多尺度性质有关。

给出图4.1所示的原始信号v=（6,4,5,1），我们计算平均值和半差，得到图4.2。我们得到系数c3=1和c4=2。然后，

|  |  |  |  |
| --- | --- | --- | --- |
| 网络错误 |  | | |
|  | 网络错误 | 网络错误 |
| 网络错误 |
|  |

图4.1：原始信号V

我们再次计算平均值和半差，得到图4.3。我们得到系数c1=4和c2=1。请注意，原始信号V可以从图4.2中的两个信号重新连接，图4.2左侧的信号可以从图4.3中的两个信号重新构造。

|  |  |  |  |
| --- | --- | --- | --- |
| 网络错误 | 网络错误 |  |  |
| 网络错误 | 网络错误 |

二

一

2

图4.2：第一个平均值和上半年的差异

|  |  |  |  |
| --- | --- | --- | --- |
| 网络错误 | 网络错误 | 网络错误 | 网络错误 |

1 1

### -1-1个

图4.3：第二个平均值和下半年差异

该方法可推广到任意长度2n的信号，前一种情况对应n=2。让我们考虑n=3的情况。HAAR基础（w1、w2、w3、w4、w5、w6、w7、w8）是

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 网络错误 |  |  |  |  |  |  |  |
| 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误   1. 网络错误 2. 网络错误   网络错误 |

这个矩阵的列是正交的，很容易看出

w−1=diag（1/8,1/8,1/4,1/4,1/2,1/2,1/2,1/2）w>。

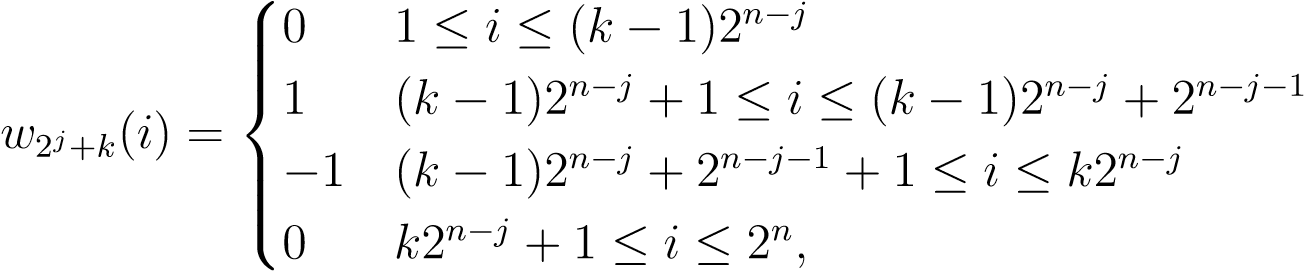
一种模式开始出现。看起来第二个haar基向量w2是所有其他基向量的“母亲”，除了第一个，它的目的是执行平均。实际上，一般来说，

，

其他的haar基向量是通过“缩放和移动过程”获得的。从w2开始，缩放过程生成向量。

W3，W5，W9，…，W2J+1，…，W2N−1+1，

这样，W2J+1+1通过形成两个连续的块（1和−1是W2J+1中块大小的一半），并将所有其他条目设置为零，从W2J+1获得。观察W2J+1有2J块2N−J元素。移动过程包括通过从左侧插入一个（k−1）2n−j零块（0≤j≤n−1和1≤k≤2j）将w2j+1中的1和−1块向右移动。因此，我们得出w2j+k的以下公式：



当0≤j≤n−1和1≤k≤2j.当然

.

如果我们通过让k从0到2j−1变化，并使用索引j而不是2j来稍微改变我们的索引，上面的公式看起来会更好一些。

定义4.5.维2n的haar基向量表示为

，

哪里

0 1≤i≤k2n−j

\_\_

j 1 k2n−j+1≤i≤k2n−j+2n−j−1 hk（i）=−k2n−j+2n−j−1+1≤i≤（k+1）2n−j

1\_

0（k+1）2n−j+1≤i≤2n，

当0≤J≤N−1和0≤K≤2J−1时。列为向量的2n×2n矩阵

，

（按这个顺序），称为维2n的haar矩阵，用wn表示。

结果表明，如果我们将向量u=（u1，…，um）解释为区间[0,1]上的分段线性函数，有一种更好地理解这些公式的方法。

定义4.6.给定一个向量u=（u1，…，um），分段线性函数plf（u）被定义为plf（u）（x）=ui，

换句话说，函数plf（u）在间隔[0,1/m]上具有值u1，在[1/m，2/m]上具有值u2，在间隔[m−1/m，1]上具有值um。

0

0.1

0.2

0.3

0.4

0.5

0.6

0.7

0.8

0.9

1

−

2

−

1

0

1

2

3

4

5

6

7

图4.4：分段线性函数plf（u）

例如，与向量关联的分段线性函数

U=（2.4、2.2、2.15、2.05、6.8、2.8、−1.1、−1.3）

如图4.4所示。

然后，每个基向量对应于函数

.

特别是，对于所有n，haar基向量

h00=w2=（1，…，1，−1，…，−1）

|新西兰2

得出相同的分段线性函数ψ，由

|  |  |
| --- | --- |
|  | 网络错误 |

如图4.5所示。那么，很容易看出这是由简单表达式ψk j（x）=ψ（2jx−k），0≤j≤n−1，0≤k≤2j−1给出的。

J

上面的公式表明，ψk是由ψ通过缩放和移动得到的。

定义4.7.函数）是一个分段线性函数，在[0,1]上有一个常量值1，函数）与haar小波一起被称为haar小波。

我们可以使用更快的算法，使用平均和差分，而不是使用w−1在haar基础上将向量u转换为系数的向量c，并使用矩阵w从haar系数c重建向量u。

1

1

−

1

0

图4.5:Haar小波ψ

如果c是维2n的Haar系数的向量，我们计算向量u0，u1，…，un的序列如下：

U0= C

UJ+1=UJ

UJ+1（2I−1）=UJ（I）+UJ（2J+I）UJ+1（2I）=UJ（I）−UJ（2J+I）、

对于j=0，…，n−1和i=1，…，2j，重构向量（信号）为u=un。

如果u是维2n的向量，我们计算向量cn、cn-1、…、c0的序列如下：

CN= U

Cj=Cj+1

Cj（i）=（Cj+1（2i−1）+Cj+1（2i））/2 Cj（2j+i）=（Cj+1（2i−1）−Cj+1（2i））/2，

对于j=n−1，…，0和i=1，…，2j，Haar基上的向量是c=c0。

我们把它作为一个练习，用两个变量u和c在matlab中实现上述程序，并通过迭代构建2j，这里是一个将n=3的向量转换为其haar系数的例子。

给定序列U=（31,29,23,17、-6、-8、-2、-4），我们得到序列

c3=（31、29、23、17、-6、-8、-2、-4）c2=（30、20、-7、-3、1、3）c1=（25、-5、5、-2、1、3、1、1）c0=（10、15、5、-2、1、3、1、1）

所以c=（10,15,5，−2,1,3,1,1）。相反，给定c=（10,15,5、-2,1,3,1,1），我们得到序列

U0=（10,15,5、-2,1,3,1,1）U1=（25、-5,5、-2,1,3,1,1）U2=（30,20、-7、-3,1,3,1,1）U3=（31,29,23,17、-6、-8、-2、-4），

它返回u=（31、29、23、17、-6、-8、-2、-4）。

有另一种递归的方法来构造维数2n的haar矩阵wn，这使得wn的列为什么是成对正交的，以及为什么上述算法确实是正确的（似乎没有人证明！）.如果我们将wn分解为两个2n×2n-1矩阵，那么包含wn最后2n-1列的第二个矩阵具有非常简单的结构：它由向量组成。

（1、−1,0，…，0）

|新西兰2

以及2n−1−1移位的副本，如下图所示，n=3：

10 0 0\_

−10 0 0\_

0 1 0 0\_

γ

0−1 0 0

γ

0 0 1 0。

 

0 0−1 0\_

 

0 0 1\_

0 0 0 0−1

注意，在我们的例子中，这个矩阵可以从单位矩阵I2n-1中获得。

，

通过形成2n×2n−1矩阵，用列向量替换每个1得到矩阵

以及每一个零点的列向量

现在，wn的前半部分，即由wn的前2n-1列组成的矩阵，可以通过形成2n×2n-1矩阵从wn-1中获得，该矩阵通过用列向量替换每个1获得。

每个−1的列向量

，

以及每一个零点的列向量

对于n=3，w3的前半部分是矩阵

1 1 0\_

1 1 0\_

1 1−1 0\_

1 1−1 0\_

1−1 0 1

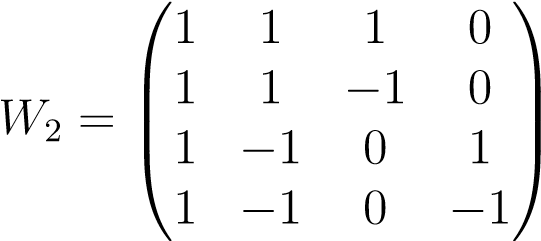
 

1−1 0 1\_

 

1−1 0−1 1−1 0−1

这确实是从



使用我们刚才描述的过程。

这些矩阵操作可以通过对被称为Kronecker产品的矩阵进行产品操作来方便地描述。

定义4.8.给定m×n矩阵a=（aij）和p×q矩阵b=（bij），a和b的Kronecker积（或张量积）a b是mp×nq矩阵。

|  |  |  |  |
| --- | --- | --- | --- |
|  | 网络错误 | 网络错误 |  |

A B=…····································

AM1B AM2B···AMNB

可以看出是关联的，并且

（a b）（c d）=ac bd

（a b）>=a>b>，

只要AC和BD定义良好。然后，立即通过以下简洁的递归方程验证Wn：

，

用w0=（1）。如果我们让

对于n≥1，

那么，不难得到方程的严格证明。

，对于所有n≥1。

上述方程清楚地证明了wn的柱是成对正交的。

观察右块（尺寸为2n×2n−1）清楚地显示了如何在n−1步后将向量c后半部分中的细节系数相加并减去到部分重建向量后半部分中的条目中。

HAAR基础的一个重要和有吸引力的特点是它提供了信号的多分辨率分析。事实上，给定信号u，如果c=（c1，…，c2n）是其haar系数的矢量，低指数的系数给出关于u的粗略信息，高指数的系数表示精细信息。例如，如果u是一个对应于管弦乐队演奏的莫扎特协奏曲的音频信号，c1对应于“背景噪声”，c2对应于低音，c3对应于第一大提琴，c4对应于第二大提琴，c5，c6，c7，c7对应于中提琴，那么小提琴等。小波的多分辨率特性可以被用来压缩信号，也就是说，用更少的系数来表示它。下面是一个例子。考虑信号

U=（2.4、2.2、2.15、2.05、6.8、2.8、−1.1、−1.3），

其haar变换为c=（2,0.2,0.1,3,0.1,0.05,2,0.1）。

与u和c对应的分段线性曲线如图4.6所示。由于c中的一些系数很小（小于或等于0.2），我们可以用0替换它们来压缩c。我们得到c2=（2,0,0,3,0,0,2,0），

重建信号是

U2=（2,2,2,2,7,3、−1、−1）。

图4.7显示了对应于u2和c2的分段线性曲线。

−

1

0

1

2

3

4

5

6

7

0.5

1

1.5

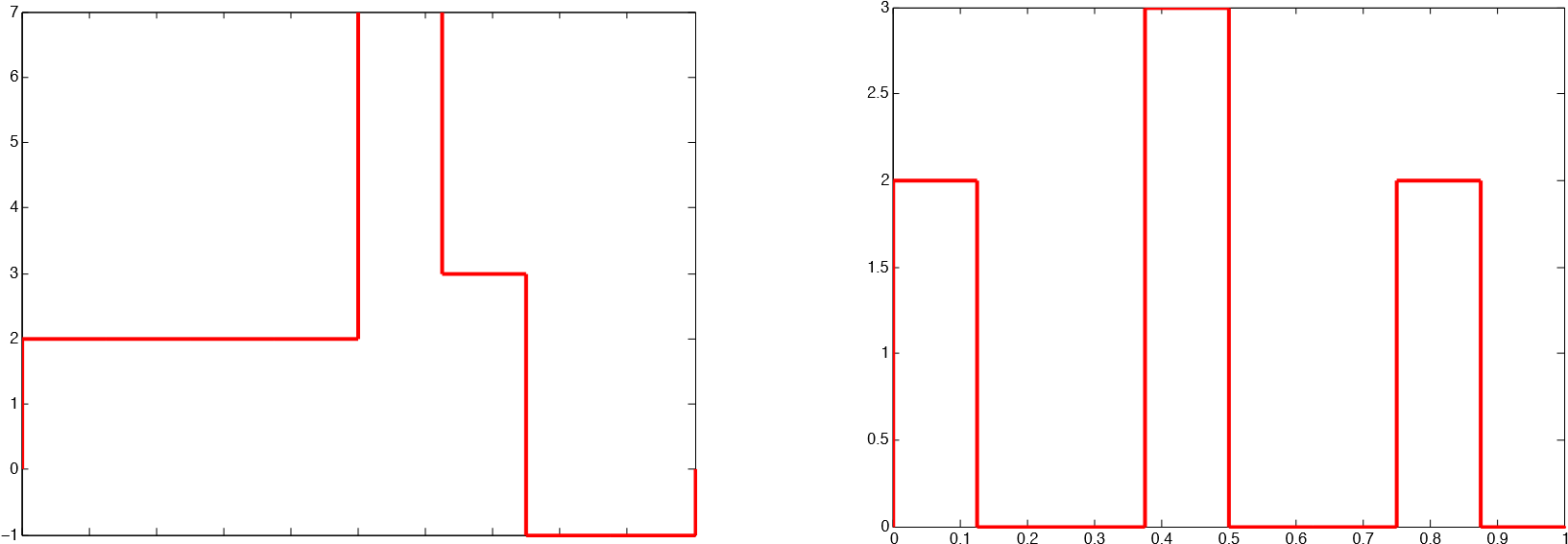
2

2.5

3

-20 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1 00 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

图4.6：信号及其haar变换



0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

图4.7：压缩信号及其压缩haar变换

HAAR小波的一个有趣的（有趣的）应用是音频信号的压缩。事实证明，如果您的类型在matlab中加载handel，音频文件将加载到y表示的向量中，如果您键入sound（y），计算机将播放这段音乐。你可以把y转换成haar系数c的向量，y的长度是

−

0.6

−

0.4

−

0.2

0

0.2

0.4

0.6

0.8

−

0.6

−

0.4

−

0.2

0

0.2

0.4

0.6

−0.80 1 2 3 4 5 6 7−0.80 1 2 3 4 5 6 7 x 104 x 104

图4.8：信号“handel”及其haar变换

73113，所以首先对y的尾部进行调整，得到长度为65536=216的向量。图4.8显示了与y和c对应的信号图。然后，运行一个程序，将绝对值小于0.05的C的所有系数设置为零。这将37272系数设置为0。结果向量c2被转换成信号y2。Y2和C2对应的信号图如图4.9所示。当你输入声音（Y2）时，你会发现音乐

−

1

−

0.8

−

0.6

−

0.4

−

0.2

0

0.2

0.4

0.6

0.8

1

−

0.8

−

0.6

−

0.4

−

0.2

0

0.2

0.4

0.6

0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7 x 104 x 104

图4.9：压缩信号“handel”及其haar变换

虽然听起来不那么清脆，但与原作差别不大。你应该和其他大于或小于0.05的数字一起玩。你应该听到当你输入声音（c）时会发生什么。它播放与y的haar变换c相对应的音乐，非常有趣。

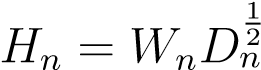
HAAR变换的另一个优点是它可以立即推广到矩阵（甚至矩形）而不需要任何额外的努力！这样可以压缩数字图像。但首先，我们讨论了HAAR系数的归一化问题。正如我们之前所观察到的，haar基向量的2n×2n矩阵wn具有正交列，但其列没有单位长度。因此，不是wn的倒数，而是矩阵



DN=诊断

20 21 22 2N−1

定义4.9.正交矩阵



其列是标准化的haar基向量，

=迪亚格

称为归一化Haar变换矩阵。给定一个向量（信号）u，我们称之为u的归一化haar系数。

因为hn是正交的。

然后，一个反射的时刻表明，我们必须稍微修改计算算法和hnc，如下所示：当计算ujs的序列时，使用

，

当计算cjs的序列时，使用

.

注意，现在情况更加对称，以√2除法为代价。然而，对于长向量，这些算法在数值上更稳定。

备注：部分作者（如Stollnitz、Derose和Salesin[163]）将c重新缩放1/√2n

u乘以√2n，这是因为基函数ψkj的范数不等于1（内积下）。归一化基函数是函数

.

现在让我们解释一下HAAR转换的二维版本。我们使用矩阵wn来描述版本，使用hn的方法是相同的（除了，但这不适用于wn-1）。给定一个2 m×2n矩阵A，我们可以首先使用haar变换wn-1将a的行转换为其haar系数，获得矩阵b，然后使用矩阵wm-1将b的列转换为其haar系数。因为列和行是在第一步交换的，

B=A（wn-1）>，

在第二步c=wm−1b，因此，我们

.

在另一个方向上，给定一个haar系数的矩阵c，我们首先将wm应用于c的列，得到b，然后再应用于b的行，从而重建矩阵a（图像）。

.

当然，我们实际上不需要转换wm和wn，也不需要执行矩阵乘法。我们只需要使用平均和差分的算法。下面是一个例子。

如果数据矩阵（图像）是8×8矩阵

|  |  |  |  |
| --- | --- | --- | --- |
| 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 |
| 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 |

如我们所见，C的条目数比A多0个；它是A的压缩版本。我们可以通过设置为0进一步压缩C，所有绝对值的条目数最多为0.5。然后，我们得到

|  |  |  |
| --- | --- | --- |
| 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 |
| 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 |
| 网络错误 | |

结果表明，matlab有一个很好的命令image（x）（也就是imagesc（x），它通常做得更好），它显示矩阵x有一个图像，其中每个条目显示为一个小正方形，其灰度与该条目的数值成比例（如果值更高，则更轻，如果该值接近零，则变暗；负值被视为零）。A和C对应的图像如图4.10所示。这个

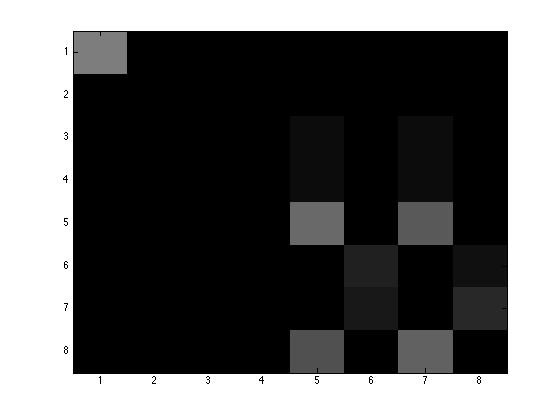
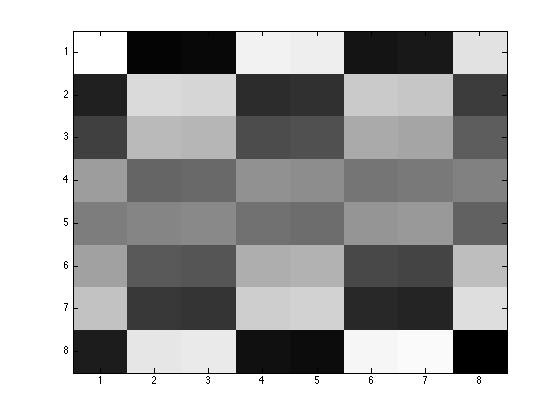


图4.10：图像及其haar变换

A2和C2对应的压缩图像如图4.11所示。压缩的

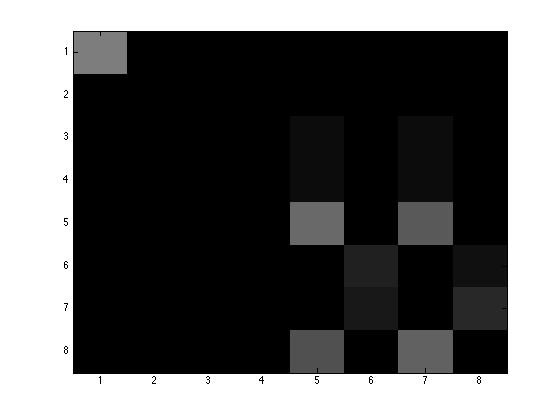
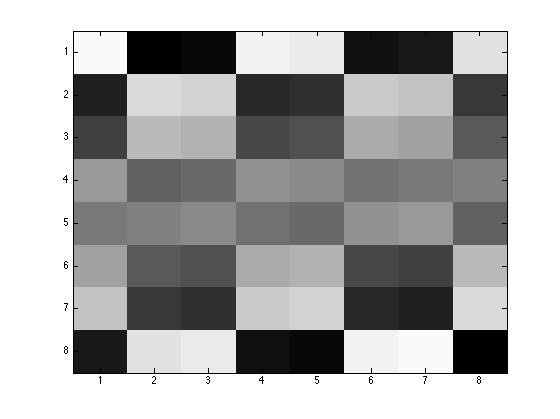


图4.11：压缩图像及其haar变换

版本似乎与原始版本不可区分！

如果我们使用归一化矩阵hm和hn，那么图像矩阵a及其归一化haar变换c的相关方程是

c=hm>ahn a=hmchn>。

HAAR转换还可以用于在互联网上逐步发送大图像。实际上，我们可以从最粗的系数（第一列自上而下，然后第二列等）开始发送矩阵c的haar系数，在接收端，我们可以在收到足够的数据后立即开始重建图像。

注意，我们可以执行部分编码（和解码），而不是对每一行和每一列执行所有轮的平均和差分。例如，我们可以对每一行和每一列执行一轮平均和差分。结果是一个包含四个子图像的图像，其中左上角的四分之一是原始图像的较粗版本，其余的（由三个部分组成）包含最精细的细节系数。我们还可以执行两轮平均和差分，或三轮等。该过程如图4.12所示。执行一轮、两轮、三轮和九轮平均的结果如图4.13所示。由于我们的图像大小为512×512，九轮平均将生成haar转换，显示为右下角的图像。原来的图像已经完全消失了！我们将它作为一个有趣的练习来修改涉及平均和差分的算法，以执行k轮平均/差分。重建算法有点复杂。

在Stollnitz、Derose和Salesin[163]中可以找到小波及其在图像处理和计算机图形中的用途的一个很好且容易访问的描述。非常详细的

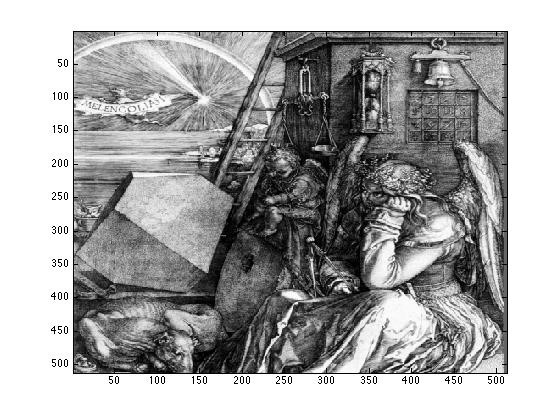


图4.12：杜勒的原始图纸

在Strang and and Nguyen[167]中给出了说明，但本书在信号处理方面有相当多的背景。

我们可以很容易地找到线性映射的2n×2n=22n向量wij（2n×2n矩阵）的基础，该线性映射从其haar系数重建图像，在任何haar系数的矩阵c的意义上，图像矩阵a由下式给出：

2N 2N

A=XXCIJWIJ。

I=1 J=1

实际上，矩阵wij是由所谓的外积给出的。

wij=wi（wj）>。

同样，对于二维Haar变换，有一个2n×2n=22n向量hij（2n×2n矩阵）的基础，在任何矩阵a中，其Haar系数的矩阵c由

2N 2N

C=XXaijhij.

I=1 J=1

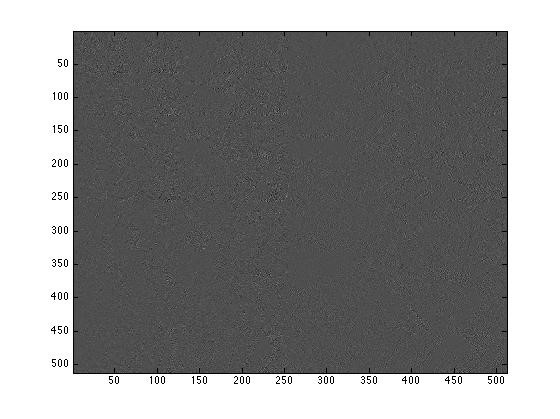
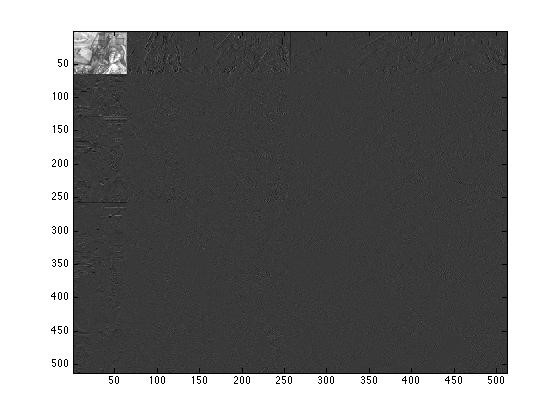
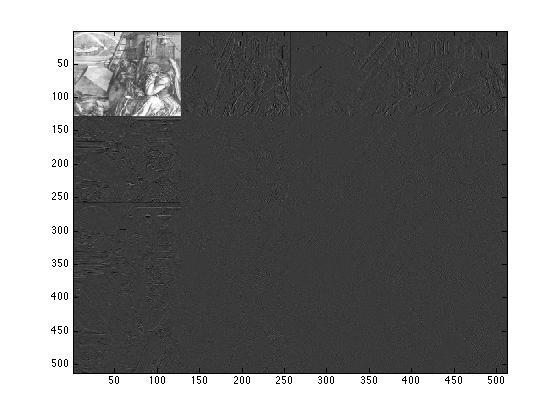
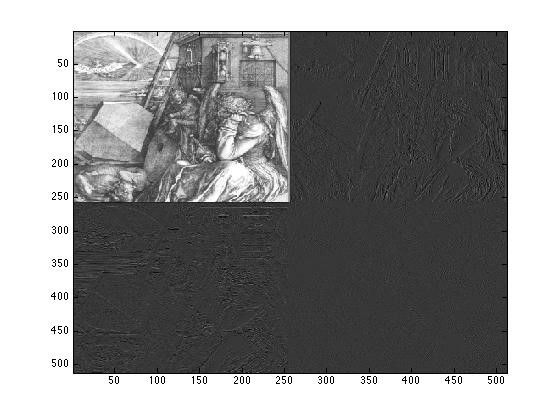


图4.13:1轮、2轮、3轮和9轮平均后的haar变换

4.4。基的变化对矩阵的影响

如果w-1的列是，那么

hij=wi0（wj0）>。

我们将它作为练习来计算n=2的基（wij）和（hij），并使用命令imagesc显示相应的图像。

## 4.4基的变化对矩阵的影响

下面的命题描述了基的变化对线性映射表示的影响。

提案4.4.设e和f为向量空间，设u=（u1，…，un）和

是e的两个基，让v=（v1，…，vm）是f的两个基。

p=p u0，u是基矩阵从u到u0的变化，Q=p v0，v是基矩阵从v到v0的变化。对于任何线性映射f:e→f，设m（f）=m u，v（f）为与f w.r.t相关联的矩阵。基U和v，设m0（f）=m u0，v0（f）为与f w.r.t相关联的矩阵。基U0和v0。我们有

m0（f）=q−1m（f）p，

或者更明确地说

.

证据。因为f:e→f可以写成f=idf f ide，因为p是ide的矩阵

w.r.t.基（）和（u1，…，un）和q−1是idf w.r.t的矩阵。根据命题4.2，基（v1，…，vm）和（），我们得到m0（f）=q−1m（f）p。

作为推论，我们得到以下结果。

推论4.5。设e为向量空间，设e的两个基，设p=p u0，u为基矩阵从u到u0的变化。对于任何线性映射f:e→e，设m（f）=m u（f）为与f w.r.t.相关联的矩阵，以u为基础，并设

m0（f）=mu0（f）是与f w.r.t.相关的矩阵，即基U0。我们有

m0（f）=p−1m（f）p，

或者更明确地说，

.

例4.3.设e=r2，u=（e1，e2），其中e1=（1,0）和e2=（0,1）是规范基向量，设v=（v1，v2）=（e1，e1−e2），并设

.

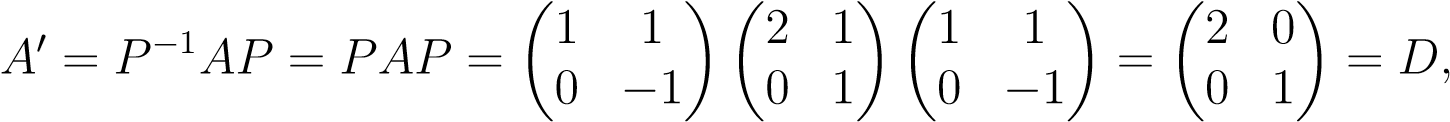
基矩阵p=p v，u从u变为v的变化是

，

我们检查一下

P−1=P.

因此，在基V中，表示a定义的线性映射f的矩阵是



对角矩阵。在基础V中，很清楚f的作用是什么：它在v1方向上是一个系数为2的拉伸，它是在v2方向上的标识。观察v1和v2不是正交的。

我们对矩阵A进行了对角化，对角项2和1是a（和f）的特征值，v1和v2是对应的特征向量。稍后我们将回到特征值和特征向量。

上面的例子表明，同一个线性映射可以用不同的矩阵表示。这建议作出以下定义：

定义4.10.两个n×n矩阵a和b是相似的，如果有一些可逆矩阵p，那么

B=P−1安培。

很容易看出相似性是等价关系。根据我们之前的考虑，两个n×n矩阵a和b是相似的iff，它们代表两个不同基的相同线性映射。以下令人惊讶的事实可以证明：每个方阵A与其转置A>相似。证明需要先进的概念（约旦形式，或相似不变量）。

如果u=（u1，…，un）和v=（v1，…，vn）是e的两个基，则基矩阵的变化

|  |  |  |  |
| --- | --- | --- | --- |
| 网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 |

4.4。基的变化对矩阵的影响

从（u1，…，un）到（v1，…，vn）是一个矩阵，其jth列由基础上的vj坐标组成（u1，…，un），这意味着

.

扩展矩阵表示法并将向量表示为

矩阵的乘积乘以向量，即

，

但请注意，所涉及的矩阵不是p，而是它的转置p>。

这个观察结果如下：如果u=（u1，…，un）和v=（v1，…，vn）是e和if的两个基

，

也就是说，

对于任何向量w∈e，如果

然后

xn yn

如此

y1 x1

…=（a>）-1…。YN XN

很容易看出（a>）-1=（a-1）>。另外，如果u=（u1，…，un），v=（v1，…，vn），w=（w1，…，wn）是e的三个基，如果基矩阵从u变为v是

p=p v，u，基矩阵从v到w的变化是q=pw，v，那么

v1 U1

…=P>…，VN UN

所以

w1 v1

…=Q>\_…，WN VN

W1 U1 U1

…=Q>P>…=（PQ）>…，WN UN联合国

也就是说，基矩阵pw，u从u到w的变化就是pq。这证明了pw，u=pv，upw，v。

尽管矩阵是必不可少的，因为它们是线性代数应用中的主要工具，但人们不应忘记以下事实：

线性映射更为基本，因为它们是不依赖于基的选择的内在对象。

因此，我们建议读者尝试用线性映射来思考，而不是把所有的东西都简化成矩阵。

根据我们的经验，这在证明线性映射和矩阵的结果时特别有效，其中涉及线性映射的证明通常更“概念性”。这些证明通常更通用，因为它们不依赖于维度是有限的事实。此外，不把矩阵分解看作纯粹的代数运算，而是把它看作几何分解。这就是SVD的情况，用几何术语来说，每个线性映射都可以作为一个旋转因子，然后沿着正交轴重新缩放，然后再进行另一个旋转。

毕竟，A

矩阵是线性映射的表示。

矩阵的大多数分解都反映了这样一个事实，即只要选择合适的基（或基），线性映射就是由具有特殊形状的矩阵表示的。问题是找到这样的基础。

然而，对于初学者来说，矩阵有着不可抗拒的吸引力，我们承认，要达到处理线性映射变得更加自然的程度，需要一定的实践。我们仍然推荐它！例如，尝试将用矩阵表示的结果转换成用线性映射表示的结果。每当我们尝试这个练习时，我们都学到了一些东西。

同时，要时刻记住

线性地图本质上是几何的；它们作用于空间。

4.5。总结

## 4.5总结

本章的主要概念和结果如下：

* 用矩阵表示线性映射。
* 线性映射的向量空间homk（e，f）。
* 同构（命题4.2）。矩阵表示映射m:hom（e，f）→mn，p和表示
* Haar基向量和Haar小波的一瞥。
* 矩阵的克罗内克积（或张量积）。
* 基矩阵和命题4.4的变化。

118第4章。矩阵和线性映射

第五章

# 直接和

## 5.1总额、直接总额、直接产品

有一些从旧的向量空间形成新的向量空间的有用方法，特别是直接积和直接和。关于直接和，有一个微妙的点，即如果我们试图用笛卡尔积e×f来定义两个向量空间的直接和e`f，由于e`f=f`e的元素，我们没有得到正确的概念。因此，我们想把e×f的元素按顺序排列。对，但我们需要未重新排列的元素对。通过考虑一个家庭的直接总和是可能的。

（ei）i 1,2，更一般地说是家族（ei）i i。为了简单起见，我们首先考虑i 1,2的情况。

定义5.1。`在两个向量空间的一个族（2（或余积）eof the family 1）i∈1,2中，我们将（ei）i∈1,2定义为集（外部）

直接和e1

e1 ae2=h1，ui，h2，vi u∈e1，v∈e2，

添加

h1，u1i，h2，v1i+h1，u2i，h2，v2i=h1，u1+u2i，h2，v1+v2i，

和标量乘法

λh1，ui，h2，vi=h1，λui，h2，λvi。

我们将in1:e1→e1`e2和in2:e2→e1`e2中的注入定义为线性映射，这样in1（u）=h1，ui，h2,0i，

且in2（v）=h1,0i，h2，vi。

一百一十九

注意

e2 ae1=h2，vi，h1，ui v∈e2，u∈e1=e1 ae2.

因此，e1`e2的每个成员h1、ui、h2、vi可以被视为一个无序对，由两个向量u和v组成，分别用索引1和2标记。

注：事实上，e1`e2只是家族（ei）i 1,2 ei的产物。

这不能与笛卡尔积e1×e2混淆。向量空间e1×e2是所有有序对hu，vi的集合，其中u∈e1，v∈e2，加上和乘上一个标量定义如下：

hu1，v1i+h u2，v2i=hu1+u2，v1+v2i，λhu，vi=hλu，λvi.

qi∈1,2 ei和e1×e2之间有一个双射，但正如我们刚才所看到的，qi∈1,2 ei的元素是一定的集合。也可以定义任意数量向量空间的积e1×····×en。我们很快就会做到。

以下属性保留。

提案5.1.给定任意两个向量空间e1和e2，集合e1`e2是向量空间。对于每对线性映射，f:e1→g和g:e2→g，都有一个唯一的线性映射，f+g:e1`e2→g，这样（f+g）in1=f和（f+g）in2=g，如下图所示：

EIN1E‘1EPP2PPPPFPPPNPN’/7 g f+g

2欧nnnn

在nnnng

神经网络神经网络

E2

证据。定义

（f+g）（h1，ui，h2，vi）=f（u）+g（v）

对于每个u∈e1和v∈e2。立即验证F+G是具有所需性质的唯一线性映射。

我们已经注意到e1`e2与e1×e2是双射的。如果我们定义投影π1:e1`e2→e1和π2:e1`e2→e2，这样

π1（h1，ui，h2，vi）=u，

π2（h1，ui，h2，vi）=v，

我们有以下建议。

提案5.2.对于任意两个向量空间e1和e2，对于每对线性映射，f:d→e1和g:d→e2，都有一个唯一的线性映射，f×g:d→e1`e2，这样π1（f×g）=f和π2（f×g）=g，如下图所示：

FN nnnnnn 7 Eo 1π1

神经网络

D PNPNPNPNPNPFGP×PPGPPP/PPEP1（E'π22E2

证据。定义

（f×g）（w）=h1，f（w）i，h2，g（w）i，

对于每一个w∈d，立即证明f×g是具有所需性质的唯一线性映射。

注：有限族的直和与积同构是线性代数的一个特点。然而，对于无限族的积和和来说，这不再是正确的。

当u，v是向量空间e的子空间时，让i1:u→e和i2:v→e作为包含映射，如果u'v在命题5.1给出的i1+i2映射下与e同构，我们说e是u和v的直接和，我们写e=u'v（略带滥用符号，因为e和u'v a只是同构的）。定义e的任意子空间的和u1+············································

定义5.2.如果p≥2个向量空间e1，…，ep，则积f=e1×···································

（u1，…，向上）+（v1，…，vp）=（u1+v1，…，向上+vp）λ（u1，…，向上）=（λu1，…，λ向上）

对于所有的ui，vi∈ei和allλ∈r，e1×·····×ep的零向量是p-元组。

，

其中，第i个零是ei的零向量。

通过上述加法和乘法，向量空间f=e1×······×ep称为向量空间e1，…，ep的直积。

作为一种特殊情况，当e1=····=ep=r时，我们又找到了向量空间f=rp，投影图pri:e1×·····×ep→ei由

pri（u1，…，向上）=ui

显然是线性的。同样，地图ini:ei→e1×····×ep由

ini（ui）=（0，…，0，ui，0，…，0）

是内射的和线性的。如果dim（ei）=ni，如果（）是i=1，…，p的ei基础，那么很容易看出n1+····+np向量

（e11,0，…，0），…，（e1n1,0，…，0），

………

（0，…，0，ei1，0，…，0），…，（0，…，0，eini，0，…，0），

………

（0，…，0，ep1），…，（0，…，0，epnp）

构成e1×······×ep的基础，依此类推。

尺寸（e1×·····×ep）=dim（e1）+····+dim（ep）。

现在让我们考虑向量空间e和p的子空间u1，…，向上到e。我们有一个地图。

A:U1×·····×向上→E

由a（u1，…，up）=u1+·····+up给出，

当ui∈ui表示i=1，…，p时，很明显这张地图是线性的，因此它的图像是e的一个子空间，表示为

U1+·····+向上

称为子空间U1的和，…，向上。根据定义，

U1+·······+上·····································

并立即证明U1+·····································这也意味着U1+······+向上不依赖于因子ui的顺序；特别是，

u1+u2=u2+u1。

如果图A是内射的，那么根据命题3.12，我们得到了

每个0都是e的零向量，这意味着如果ui∈ui代表i=1，…，p和if

U1+·····+向上=0，

那么（u1，…，up）=（0，…，0），也就是说，u1=0，…，up=0。在这种情况下，每个u∈u1+······+up都有一个独特的表达式作为和

U=U1+·····+向上，

对于ui∈ui，对于i=1，…，p.确实，如果

u=v1+····+vp=w1+····+wp，

对于vi，wi∈ui，对于i=1，…，p，那么我们有

w1−v1+····+wp−vp=0，

既然vi，wi∈ui，每个ui都是一个子空间，wi−vi∈ui。a的注入性意味着wi-vi=0，也就是说，wi=vi代表i=1，…，p，这表明u的分解是唯一的。

很明显，任何p非零向量u1，…，上到ui∈ui都是线性无关的。为此，假设λ1u1+····+λpup=0

对于某些λi∈r，由于ui∈ui和ui是一个子空间，因此，对于i=1，…，p，a的内射性意味着λiui=0，对于i=1，…，p。由于ui=06，我们必须得到λi=0，对于i=1，…，p；也就是说，u1，…，up with ui∈ui和ui=06是线性无关的。

如果a是注射剂，那么当i=6 j时，我们必须有ui uj=（0）。但是，如果p≥3，这种情况通常是不够的。例如，如果e=r2和u1，用e1=（1,0）表示的线，u2是用d=（1,1）表示的线，u3是用e2=（0,1）表示的线，那么u1 u2=u1 u3=u2（0,0），但是u1+u2=u1+u3=u2+u3=r2，所以u1+u2+u3不是直接和。例如，d用两种不同的方式表示为

d=（1,1）=（1,0）+（0,1）=e1+e2。

定义5.3.对于任意向量空间e和任意p≥2个子空间u1，…，向上到e，如果上面定义的映射a是内射的，则该和u1+·····+向上称为直接和，并用

U1·····向上。空间e是子空间ui的直接和，如果

E=U1····向上。

如求和，u1 u2=u2 u1。观察到当图A为内射时，则为U1×上×上与U1×上的线性同构。区别在于，即使空间ui不被假定为某个公共空间的子空间，也定义了u1×······×up。

如果e是一个直接和e=u1········up，由于任何p非零向量u1，…，up与ui∈ui是线性无关的，如果我们在uj中选取一个基（uk）k∈i j，对于j=1，…，p，那么（ui）i∈i与i=i1·······ip是e的基，直观地，e被分成p独立的子空间。

相反，给定e的基（ui）i∈i，如果我们将索引集i划分为i=i1···ip，则每个子族（uk）k∈ij跨越e的一些子空间uj，并立即证明我们有一个直接和

E=U1····向上。

设f:e→e为线性映射。如果f（uj）uj，我们说uj在f下是不变的，假设e是有限维的，直接和e=u1·································································在f下，每个基向量uk的k∈ij的图像f（uk）都属于uj，因此表示f的矩阵a在基（ui）i∈i上是一个块对角矩阵的形式

，

对于每个块aj，一个dj×dj矩阵，dj=dim（uj），所有其他条目等于0。如果j=1时dj=1，…，p，则矩阵a是对角矩阵。

每个ui到e都有自然注入，由ini:ui→e表示。

现在，如果p=2，很容易确定图a:u1×u2→e的核心，我们有

a（u1，u2）=u1+u2=0 iff u1=−u2，u1∈u1，u2∈u2，

这意味着

kera=（u，−u）u∈u1 u2。

现在，u1 u2是e的一个子空间，线性映射u→7（u，−u）显然是u1 u2和kera之间的同构，所以kera与u1 u2是同构的。因此，我们得到以下结果：

提案5.3.给定任意向量空间e和任意两个子空间u1和u2，和u1+u2是直接和iff u1 u2=（0）。

直接和概念的一个有趣的例子是将一个方阵分解成它的对称部分和它的斜对称部分。回想一下，n×n矩阵a∈mn在a>=a时是对称的，在a>=-a时是斜对称的。

s（n）=a∈mn a>=a和skew（n）=a∈mn a>=-a

是mn的子空间，s（n）skew（n）=（0）。观察任意矩阵a∈mn，矩阵h（a）=（a+a>）/2是对称的，矩阵s（a）=（a−a>）/2是偏对称的。自从

，

我们看到mn=s（n）+skew（n），由于s（n）skew（n）=0，我们得到了直接和

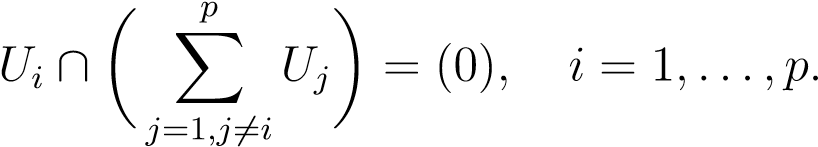
mn=s（n）倾斜（n）。

注：斜对称矩阵的向量空间斜（n）也用so（n）表示。它是群so（n）的李代数。

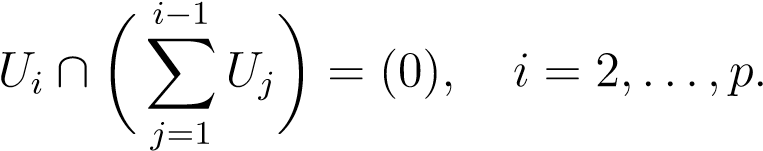
命题5.3可以以符号为代价推广到任何p≥2个子空间。下列命题的证明留作练习。

提案5.4.给定任意向量空间e和任意p≥2个子空间u1，…，向上，以下属性等效：

1. u1+········+上的和是直接和。
2. 我们有



1. 我们有



因为同构

U1×上×上≈U1×上，

我们有

提案5.5.如果e是任何向量空间，对于任何（有限维）子空间u1，…

在e上，我们有dim（u1·····································

如果e是直接和

E=U1····向上，

因为每个u∈e都可以用一种独特的方式写成

U=U1+·····+向上

对于i=1…，p的ui∈ui，我们可以通过以下方式定义映射πi:e→ui，称为投影：

πi（u）=πi（u1+····+向上）=ui。

很容易检查这些映射是否是线性的并且满足以下属性：

（

πjπi=πi，如果i=6 J

0，如果i=j，

π1+····+πP=ide.

例如，对于直接和

mn=s（n）歪斜（n）

S（n）上的投影由下式给出

，

斜角上的投影（n）由

A−A>π2（A）=S（A）=

二

显然，h（a）+s（a）=a，h（h（a））=h（a），s（s（a））=s（a），h（s（a））=s（h（a））。=

0。

f f=f的函数称为等幂函数。因此，投影πi是等幂的。相反，可以显示以下命题：

提案5.6.设e为向量空间。对于任意p≥2个线性映射fi:e→e，如果

（

fj fi=fi如果i 6=j

0，如果i=j，

f1+····+fp=ide，

如果我们让ui=fi（e），我们有一个直接和

E=U1····向上。

我们也有以下的命题来描述等幂线性映射，它的证明也留作练习。

提案5.7.对于每个向量空间e，如果f:e→e是一个等幂线性映射，即f f=f，那么我们有一个直接和

E=切口国际货币基金组织，

所以f是对其图像的投影，imf。

我们现在给出任意非空指数集i的一个直接和的定义。首先，让我们回忆一下一个族（ei）i∈i的积的概念。给定一个族（ei）i∈i，其积q i ei，是所有函数f:i→si∈i ei的集合，这样，f（i）∈ei，对于每个i∈

i∈i。它是选择公理的众多版本之一，如果ei=6∅对于每一个i∈i，有then qi∈iprojectionei=6∅。成员πi:qi∈i efi∈→qeii∈，定义为，i ei，通常表示为（πi（（fi）fii∈）ii∈）=i。我们现在定义∈i，我们直接求和。

定义5.4.让我是任何非空集，让（ei）i∈i是向量空间的族。家族（ei）i∈i的（外部）直接和'i∈i ei（或副产物）定义如下：

` i∈i ei由有限支持的所有f∈qi∈i ei组成，加乘一个标量定义如下：

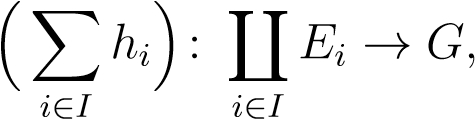
（fi）i∈i+（gi）i∈i=（fi+gi）i∈i，λ（fi）i∈i=（λfi）i∈i。

我们也有注入映射ini:ei→'i∈i ei，定义如下，ini（x）=（fi）i∈i，其中fi=x，fj=0，对于所有j∈（i−i）。

下面的命题是5.1命题的一个明显概括。

提案5.8.设我为任意非空集，设‘ei（ei，is a vector space，and for every family）i∈i为向量空间的族，设g为任意向量空间。直接和i∈i

（hi）线性映射的i∈i hi:ei→g，有一个唯一的线性映射



这样，（pi∈i hi）ini=hi，对于每一个i∈i。

评论：

1. 人们可能会想，为什么直接和‘i∈i ei由有限支持的家族而不是任意家族组成；换句话说，为什么我们不把家族的直接和（ei）i∈i定义为qi∈i ei？上面定义的加和标量乘的积空间qi∈i ei也是一个向量空间，但问题是任何线性映射bh:qi∈i ei→g都必须由

eh（（ui）∈i）=xhi（ui），

我爱我

如果我是无限的，那么右边的和是无限的，因此也没有定义！如果我是有限的，那么qi∈i ei和'i∈i ei是同构的。

1. 当e i=e时，对于所有i∈i，我们用e（i）表示‘i∈i ei’。特别是当ei=k时，对于所有i∈i，我们找到定义3.9的向量空间k（i）。

关于内射或外射线性映射，我们也有以下基本命题。

提案5.9.设e和f为向量空间，设f:e→f为线性映射。如果f:e→f是内射的，那么有一个称为收缩的投影线性映射r:f→e，这样r f=ide。如果f:e→f是主观的，那么有一个内射线性映射s:f→e称为截面，这样f s=idf。

证据。设（ui）i∈i为e的基，由于f:e→f是一个内射线性映射，由命题3.13，（f（ui））i∈i在f中是线性独立的，由定理3.5可知，f中有一个基（vj）j∈j，其中i j，其中vi=f（ui），对于所有i∈i，由命题3.13可以定义一个线性映射r:f→e。式中，r（vi）=ui，对于所有i∈i，r（vj）=w，对于所有j∈（j-i），其中w是e中的任何给定矢量，即w=0。因为r（f（ui））=ui对于所有i∈i，根据3.13号命题，我们得到r f=ide。

现在，假设f:e→f是主观的。设（vj）j∈j为f的基，因为f:e→f是可射的，对于每个vj∈f，都有一些uj∈e，这样f（uj）=vj。由于（vj）j∈j是f的基础，根据命题3.13，有一个唯一的线性映射s:f→e，因此s（vj）=uj。另外，由于f（s（vj））=vj，根据命题3.13（再次），我们必须有f s=idf。

5.9号提案的反面是显而易见的。

我们现在准备证明一个关于线性映射核的秩和维数的非常重要的结果。

## 5.2秩零定理；格拉斯曼关系

我们从以下基本命题开始。

提案5.10.设e，f和g为三个向量空间，f:e→f为一个内射线性映射，g:f→g为一个内射线性映射，并假定imf=kerg。然后，以下属性保持不变。（a）对于任意截面s:g→f of g，我们有f=kerg ims，线性映射f+s:e g→f是同构的。1

（b）对于任何收缩r:f→e of f，我们有f=imf kerr.2

弗格森

O/O/G（O/O/G）

21关于g的s:g→f节的存在，从命题5.9开始。

牵开器r的存在：f的f→e从命题5.9开始。

证据。（a）因为s:g→f是g的一部分，我们有g s=idg，并且对于每个u∈f，

G（U−S（G（U）））=G（U）−G（S（G（U）））=G（U）−G（U）=0。

因此，u−ss，然后（g（u））u∈=kers（vg），我们对一些v∈gfause=kergu+∈immss。另一方面，如果，g（u）=0，因为u∈keru g∈，

Kerg im等等，

g（u）=g（s（v））=v=0，

fbecause+s:eg gs=→fidf=g，表示同构。（b）的证明非常相似。Kerg，我们有f=uim=sf（v）=0ims。因此，但是，既然，sincef=kergf和imss，而且既然byare injective，assumption，im

F=注意，我们可以选择一个收缩kerg ims=imf ims，f是内射的，所以我们可以设置r:f→e，这样ims上的kerr im0，sinces。

弗格森

给定一个线性mapsf g e−→f−→g的序列，当imf=kerg时，我们认为该序列e−→f−→g在f处是精确的。如果除了在短精确序列上是精确的外，并且这表示为asf，f是内射的。

而G是悲观的，我们说我们有一个

0−→E−F→F−→G−→0。

命题5.10给出的短精确序列的性质通常通过说0−→e−f→f−→g g−→0是（短）分裂的精确序列来描述。

作为5.10号命题的推论，我们得到以下结果。

定理5.11.是线性映射。然后，（秩零定理）Lete与Kerefand imf fbe向量空间同构，并由此，f:e→f

dim（e）=dim（切口）+dim（imf）=dim（切口）+rk（f）。

证据。考虑

克尔

其中kerf−→i是包含图，是与Surjection关联的

f

得到一个−→f之间的同构，然后，我们将命题5.10应用于任意截面IME和切口imf，以及命题5.5，得到dim（f−→e of ef）=0 to dim（kerf）+dim（imf）。

注：线性映射f的核的维数dim（kerf）通常称为f的零。

我们现在用定理5.11得到一些重要的结果。

提案5.12。给定向量空间e，如果u和v是e的任意两个子空间，那么

调光（U）+调光（V）=调光（U+V）+调光（U V）

一个称为格拉斯曼关系的方程。

证据。回想一下，u+v是线性地图的图像

A:U×V→E

由a（u，v）=u+v给出，

我们在前面已经证明了a的kera核与u v核是同构的。根据定理

5.11，

尺寸（u×v）=dim（kera）+dim（ima）

但dim（u×v）=dim（u）+dim（v），dim（kera）=dim（u v），ima=u+v，因此

格拉斯曼关系成立。

格拉斯曼关系对于确定两个子空间在维>3的空间中是否具有非平凡交集是非常有用的。例如，很容易看出，在R5中，有子空间u和v，dim（u）=3，dim（v）=2，这样u v=（0）；例如，让u由向量（0,0,0,0,0,0），（0,1,0,0），（0,0,1,0,0）生成，v由向量（0,0,0,1,0）和（0,0,0,0,1）生成。然而，我们声称如果dim（u）=3和dim（v）=3，那么dim（u v）≥1。事实上，根据格拉斯曼的关系，我们

调光（U）+调光（V）=调光（U+V）+调光（U V）

即

3+3=6=dim（U+V）+dim（U V）

由于u+v是R5的子空间，dim（u+v）≤5，这意味着

6≤5+尺寸（U V）

即1≤dim（U V）。

作为命题5.12的另一个结果，如果u和v是维度n的向量空间中的两个超平面，那么dim（u）=n-1和dim（v）=n-1，读者应该证明

尺寸（U V）≥N−2，

所以，如果u=6v，那么

尺寸（U V）=N-2。

下面是直接从定理5.11得出的直和的特征。

提案5.13.如果u1，…，up是有限维向量空间e的任何子空间，那么dim（u1+······+up）≤dim（u1）+·····+dim（up），

和dim（u1+·····+向上）=dim（u1）+····+dim（向上）

如果统计研究所形成一个直接的总和U1·····向上。

证据。如果我们把定理5.11应用于线性映射

A:U1×······························向上

由a（u1，…，up）=u1+·····+up给出，我们得到

Dim（U1+······+向上）=Dim（U1×·······×向上）−Dim（Kera）

=dim（u1）+·····+dim（up）−dim（kera）

所以不等式如下。由于a是注射剂iff kera=（0），因此uis形成了第二个方程所持的直接和iff。

定理5.11的另一个重要推论是：

提案5.14.设e和f为两个有限维相同的向量空间dim（e）=dim（f）=n。对于每一个线性映射f:e→f，以下属性是等效的：

1. F是双主语。
2. F是悲观的。
3. F是注射剂。
4. 切口=（0）。

证据。显然，（a）意味着（b）。

如果f是主观的，那么imf=f，所以dim（imf）=n。根据定理5.11，

dim（e）=dim（切口）+dim（国际货币基金组织）

由于dim（e）=n和dim（imf）=n，我们得到dim（kerf）=0，这意味着

切口=（0），所以f是内射的（见命题3.12）。这证明（b）意味着（c）。

如果f是内射的，那么根据命题3.12，切口=（0），所以（c）意味着（d）。

最后，假设切口=（0），那么dim（切口）=0和f是内射的（根据命题3.12）。根据定理5.11，

dim（e）=dim（切口）+dim（国际货币基金组织）

因为dim（kerf）=0，我们得到

dim（imf）=dim（e）=dim（f），

证明了F也是主观性的，因此是双主观性的。这证明（d）暗示（a）并结束证明。

我们应该注意到，5.14命题在无限维上失败了。

下面的建议也很有用。

命题5.15.一个同构f:vlet→ewbe一个向量空间。如果在v和w之间，e=u\_v和e=u\_w，那么

证据。设r为v和w之间的关系，定义如下：

hv，wi∈r iff w−v∈u。

在我们声称v和r w之间，这是一个函数关系，定义了一个线性同构mf（v）=w iff hv，wu r（r是w vf的图形）。如果w f:vv u→andw

−∈v0=v，因此，

wthenw00−w vv 0∈=0−uv，即，则∈ue，且since w=w0−0u=w w∈v。因此，w，对于everyuu−，且既然vv∈ris是一个直接和，uis是泛函的。类似地，如果，andw∈i∈fwis是主观的。我们还需要验证，存在一个唯一的pairw是一个直接和，u v=（0），and u wu and=（0）hu，v wi∈，因此−∈u v 0×∈ffvusis，，

注射。自从

这样w=u+v。

线性的。如果

W−V=U

和w0−v0=u0，

其中u，u0∈u，那么，我们有

（W+W0）−（V+V0）=（U+U0）

式中，u+u0∈u。同样，如果w−v=u

式中u∈u，则我们得到λw−λv=λu，

式中，λu∈u。因此，f是线性的。

假设一个向量空间e和e的任何子空间u，命题5.15表明u的任何子空间余维的维数，我们用codim（v表示，这样e=u u）v表示它。子空间只依赖于余维1的onu称为du。我们称暗

超平面。

线性映射或矩阵的秩是一个重要的概念，无论从理论上还是实践上，它都是线性方程可解性的关键。根据定义3.16，线性映射f:e→f的秩Rk（f）是f的图像子空间imf的维数dim（imf）。

我们有以下简单的建议。

提案5.16。给定线性映射f:e→f，以下属性保持不变：

1. Rk（f）=codim（切口）。
2. Rk（f）+Dim（kerf）=Dim（e）。
3. Rk（f）≤min（dim（e），dim（f））。

证据。因为根据5.11，dim（e）=dim（kerf）+dim（imf），根据定义，rk（f）=dim（imf），我们有rk（f）=codim（kerf）。因为Rk（f）=dim（imf），（ii）来自dim（e）=dim（kerf）+dim（imf）。至于（iii），因为imf是f的一个子空间，所以我们有Rk（f）≤dim（f），并且由于Rk（f）+dim（kerf）=dim（e），我们有Rk（f）≤dim（e）。

矩阵的秩定义如下。

定义5.5.在k域上给定m×n矩阵a=（aij），矩阵a的秩Rk（a）是a线性无关列的最大数目（以km为单位视为向量）。

从命题3.6来看，矩阵a的秩是a列生成的km子空间的维数，设e和f为两个向量空间，设（u1，…，un）为e的基，（v1，…，vm）为f的基，设f:e→f为线性映射，设m（f）为其矩阵w.r.t.基（u1，…，un）和（v1，…，vm）。因为f的秩Rk（f）是由（f（u1）、…、f（un））生成的imf的维数，所以f的秩是（f（u1）、…、f（un））中线性无关向量的最大个数，等于m（f）的线性无关列的个数，因为f和km是同构的。因此，对于表示f的每个矩阵，我们有Rk（f）=Rk（m（f））。

稍后我们将看到，利用对偶性，矩阵A的秩也等于A的最大线性无关行数。

如果u是超平面，那么对于维1的某些子空间v，e=u\_v。然而，维1的子空间V是由任何非零向量V∈V生成的，因此我们用kv表示v，并写出e=u kv。显然，v/∈u。相反，让x∈e是一个向量，这样x/∈u（因此，x=0）6。我们声称e=u kx。实际上，由于u是一个超平面，对于一些v/∈u（v=0）6，我们有e=u kv。那么，x∈e可以用一种独特的方式写成x=u+λv，其中u∈u，并且由于x/∈u，我们必须有λ=06，因此，v=−λ−1u+λ−1X。由于e=u kv，这表明e=u+kx。由于x/∈u，我们得到u kx=0，因此e=u kx。证明了超平面是E的最大适配子空间H。

在第10章中，我们将看到超平面正是非零线性映射f:e→k（称为线性形式）的核心。

## 5.3总结

本章的主要概念和结果如下：

* 直积、和、直和。
* 预测。
* 基本方程

dim（e）=dim（切口）+dim（imf）=dim（切口）+rk（f）

（提案5.11）。格拉斯曼关系

调暗（U）+调暗（V）=调暗（U+V）+调暗（U V）。

* 双射线性映射f:e→f的特征。
* 矩阵的秩。

第六章

# 行列式

## 6.1排列，排列的签名

本章回顾了行列式及其在线性代数中的应用。我们从排列和排列的签名开始。接下来，我们定义多行映射和交替多行映射。行列式作为交替的多行映射引入，在单位矩阵上取1（遵循Emil Artin）。然后给出了用拉普拉斯展开式计算行列式的方法，并给出了与一般定义的联系。它展示了如何使用行列式来反转矩阵和求解（至少在理论上！）线性方程组（克莱默公式）。定义了线性映射的行列式。我们通过定义矩阵的特征多项式（和线性映射的特征多项式）和证明著名的凯莱-汉密尔顿定理得出结论，该定理指出每个矩阵都是其特征多项式的“零”（我们给出两个证明；一个是计算性的，另一个是概念性的）。

行列式可以用几种方式定义。例如，行列式可以通过向量空间的外部代数（或交替代数）以奇特的方式定义。由于Emil Artin的原因，我们将采用更为算法化的方法。无论采用哪种方法，我们都需要对有限集上的置换进行一些初步的研究。我们需要证明n个元素上的每一个置换都是置换的产物，并且所涉及的置换数目的奇偶性是置换的不变量。设[n]=1,2…，n，其中n∈n，n>0。

定义6.1.n元素的排列是双射π[n]→[n]。当n=1时，从[1]到[1]的唯一函数是常量映射：17→1。因此，我们假设n≥2。一个换位是一个置换τ[n]→[n]，这样，对于一些i<j（1≤i<j≤n），τ（i）=j，τ（j）=i，和τ（k）=k，对于所有k∈[n]−i，j。换言之，换位交换两个不同的元素i，j∈[n]。k阶的循环置换（或k-循环）是一个置换σ：【n】→【n】，因此，对于某些序列（I1、I2，…，Ik），不同元素的

[n]2≤k≤n时，

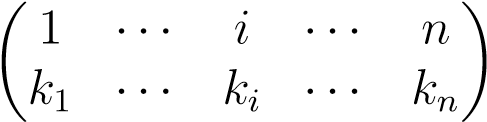
σ（i1）=i2，σ（i2）=i3，…，σ（ik−1）=ik，σ（ik）=i1，

一百三十五

且σ（j）=j，对于j∈[n]−i1，…，ik。集合I1，…，Ik称为循环置换的域，循环置换通常用（I1 I2…IK）。

如果τ是一个换位，显然，ττ=id。另外，2阶的循环置换是一个换位，对于k阶的循环置换σ，我们有σk=id。很明显，两个置换的组合是一个置换，每个置换都有一个逆，也是一个置换。因此，[n]上的排列集是一个通常表示sn的群。通过归纳，很容易看出sn组有n个！元素。我们还将使用排列（或换位）的术语产物作为排列组合的同义词。

n个元素上的置换σ，如σ（i）=ki（i=1，…，n），可用2×n数组的函数表示法表示。



被称为柯西双线符号。例如，我们的置换σ表示为

.

在计算机科学和组合数学中经常使用的一种更简洁的表示法是用图像来表示排列，即用序列来表示排列。

σ（1）σ（2）···σ（n）

作为行向量写入，不使用逗号分隔条目。以上就是所谓的单线符号。例如，在一行符号中，我们前面的置换σ表示为

2 4 3 6 5 1.

不将上述序列括在括号内的原因是避免与循环符号混淆，通常情况下，循环符号包括括号。

下面的命题说明了循环置换和换位的重要性。

提案6.1.对于每一个n≥2，对于每一个置换π[n]→[n]，有一个将[n]划分为r子集的区域，称为π的轨道，其中1≤r≤n，其中该区域中的每一个集合j要么是一个单子i，要么是形式。

j=i，π（i），π2（i），…，πri−1（i），

其中ri是最小的整数，比如πri（i）=i和2≤ri≤n。如果π不是同一性，那么它可以用一种独特的方式（按顺序）写成π=σ1…具有不相交域的循环置换的σs，其中s是具有至少两个元素的轨道数。每一个排列π：【n】→【n】都可以写成一个非空的换位组合。

6.1。排列，排列的签名

证据。考虑在[n]上定义的关系rπ，如下所示：irπj iff有一些k≥1，因此j=πk（i）。我们声称Rπ是一个等价关系。传递性是明显的。我们声称，对于每一个i∈[n]，都有一些最小r（1≤r≤n），这样πr（i）=i。

实际上，考虑以下n+1元素的序列：

嗨，π（i），π2（i），…，πn（i）i。

因为[n]只有n个不同的元素，所以存在一些h，k，其中0≤h<k≤n，这样

πh（i）=πk（i）、

由于π是双射，这意味着πk−h（i）=i，其中0≤k−h≤n。因此，我们证明了存在一些整数m≥1，这样πm（i）=i，所以存在这样一个最小的整数r。

因此，Rπ是自反的。它是对称的，因为如果j=πk（i），让r最小r≥1，这样πr（i）=i，那么

i=πkr（i）=πk（r−1）（πk（i））=πk（r−1）（j）。现在，对于每一个i∈[n]，i的等价类（轨道）是[n]的子集，或者是单子

i或一套表格

j=i，π（i），π2（i），…，πri−1（i），

式中，ri是最小整数，使得πri（i）=i和2≤ri≤n，在第二种情况下，π对j的限制导致循环置换σi和π=σ1…σs，其中s是具有至少两个元素的等价类数。

对于命题的第二部分，我们从n开始归纳，如果n=2，在[2]上正好有两个置换，置换τ交换1和2，以及同一性。然而，id2=τ2。现在，设n≥3。如果π（n）=n，因为根据诱导假设，π对[n−1]的限制可以写成转置的乘积，π本身可以写成转置的乘积。如果π（n）=k=6n，假设τ为τ（n）=k和τ（k）=n的转置，很明显，τπ保持n不变，根据诱导假设，我们得到τπ=τm……τ1表示某些转置，因此

π=ττm…τ1，

换位的产物（因为ττ=idn）。

注：当π=idn为同一置换时，我们可以认为0置换的组合是同一置换。命题6.1的第二部分表明，换位产生一组置换sn。

在写置换π作为组成π=σ1…循环置换的σs，很明显，σi的顺序并不重要，因为它们的域是不相交的。给出了一个作为转置的乘积而写的排列，我们现在证明了转置数的奇偶性是一个不变量。

定义6.2.对于每个n≥2，因为每个置换π[n]→[n]定义了R子集的一个分区，在该分区上π要么作为恒等式，要么作为循环置换，let），称为π的签名，由定义，其中r是分区中的集合数。

如果τ是交换i和j的换位，那么很明显，与τ相关的分区由n-1等价类、集合i、j和n-2单子集k组成，对于k∈[n]−i、j，因此，

提案6.2.对于每一个n≥2，对于每一个置换π，我们有：

.

因此，对于每一个转置的乘积，例如π=τm…τ1，我们有

，

这表明转置数的奇偶性是一个不变量。

证据。假设τ（i）=j和τ（j）=i，其中i<j。有两种情况，取决于i和j是在Rπ的相同等效类jl中，还是在不同的等效类中。如果我和J属于同一类，那么如果

jl=i1，…，ip，…iq，…ik，

其中ip=i和iq=j，因为

τ（π（π−1（ip）））=τ（ip）=τ（i）=j=iq

τ（π（iq−1））=τ（iq）=τ（j）=i=ip，

很明显，jl分为两个子集，其中一个子集是，因此，与τπ相关的类的数量是r+1，而j在不同的等价类jl和jm中，例如

I1，…，IP，…，IH

和

j1，…，jq，…jk，

其中ip=i，jq=j，因为

τ（π（π−1（ip）））=τ（ip）=τ（i）=j=jq

和

τ（π（π−1（jq）））=τ（jq）=τ（j）=i=ip，

6.2。交替多行地图

我们看到类jl和jm合并为一个类，因此，与τπ相关的类的数量是r−1，并且

现在，让π=τm…τ1是转位的任何产物。在命题的第一部分，我们有

，

从1开始换位。

注：当π=idn为单位置换时，由于我们一致认为0个置换的组合是单位置换，所以（（id）=+1仍然正确。从这个提议来看，这是直接的。特别是，由于π−1π=idn，我们得到

我们现在可以继续讨论行列式的定义。

## 6.2交替多行地图

首先，我们定义多行映射、对称多行映射和交替多行映射。

注：本节给出的大多数定义和结果也适用于K是交换环时，以及当我们考虑K上的模时（自由模，当需要基时）。

设e1，…，en，和f为k域上的向量空间，其中n≥1。

定义6.3.函数f:e1×…×en→f是多行映射（或n-线性映射），如果每个参数是线性的，则保持其他参数不变。更明确地说，对于每一个I，1个i i，n，对于所有的X1……E1，XI，1，1，1，1，……，对于所有x，y，Ei，所有的，

f（x1，…，Xi，1，x+y，Xi+ 1，…，xn）＝f（x1，…，Xi，1，x，Xi＋1，…，xn）

+f（x1，…，Xi，1，y，Xi+ 1，…，Xn），

f（x1，…，Xi，1，Lax x，Xi+ 1，…，Xn）＝αf（x1，…，Xi，1，x，Xi+1，…，Xn）。

当f=k时，我们称之为n-线性形式（或多行形式）。如果n≥2且e1=e2=…=e n，n-线性映射f:e×…×e→f称为对称，如果f（x1，…，x n）=f（xπ（1），…，xπ（n）），对于1，…，n上的每个置换π。n-线性映射f:e×…×e -εf被称为交错，如果f（x1，…，xn）＝0，每当Xi＝Xi＋1时，对于某些i，1的i＝n＝1（换言之，当两个相邻的参数相等）。当n=1时，线性映射被认为是对称的和交替的，我们这样做并不有害。

当n=2时，2线性映射f:e1×e2→f称为双线性映射。我们已经看到了几个双线性图的例子。乘法·：k×k→k是一个双线性映射，把k作为一个向量空间。一般来说，A环中的乘法·：A×A→A是双线性映射，将A视为一个模块。

将线性形式应用于向量的操作h−、−i:e×e→k是双线性映射。

对称双线性映射（和多线性映射）在几何（内积、二次型）和微分学（偏导数）中起着重要作用。

双线性映射是对称的，如果f（u，v）=f（v，u），对于所有u，v∈e。

交替多行映射满足以下简单但关键的特性。

提案6.3.设f:e×…×e→f为n线性交替图，n≥2。以下属性保留：

（1）f（…，Xi，Xi＋1，…）＝f（…，Xi＋1，Xi，…）

（2）

F（…，Xi，…，XJ，…）＝0，

其中Xi＝Xj，1＜i＜J＜n。

（3）f（…，Xi，…，XJ，…）＝f（…，XJ，…，XI，…）；

其中1≤i<j≤n。

（4）

F（…，Xi，…）＝f（…，X+S\* XJ，…）；

对于任何λ∈k，其中i=6 j。

证据。（1）通过应用两次多语种，我们已经

f（…，Xi+Xi+ 1，Xi+Xi+ 1，…）＝f（…，Xi，Xi，…）+f（…，Xi，Xi + 1，…）

+f（…，Xi＋1，Xi，…）+F（…，Xi + 1，Xi + 1，…）；

既然f是交替的，这就产生了

0＝f（…，Xi，Xi＋1，…）+f（…，Xi＋1，Xi，…）；

即F（…，Xi，Xi＋1，…）＝f（…，Xi＋1，Xi，…）。

1. 如果Xi＝XJ和I和J不相邻，则可以交换Xi和Xi + 1，然后交换Xi和Xi + 2等，直到XI和XJ变为相邻。由（1）

，

其中1，但F（…，Xi，XJ，…）＝0，因为Xi＝XJ，和（2）成立。

1. 从（2）开始，如（1）所示。（4）是（2）的直接后果。

6.2。交替多行地图

命题6.3现在将用于显示交替多行图的基本性质。首先，我们需要稍微扩展一下矩阵符号。设e为k上的向量空间。给定n×n矩阵a=（aij）除以k，我们可以定义一个映射l（a）：en→en，如下所示：

L（A）1（U）=A11U1+·····+A1nun，

…

L（A）N（U）=An1U1+····+阿牛，

对于所有的u1，…，un∈e，带有u=（u1，…，un）。立即证实L（a）是线性的。然后，给出两个n×n矩阵a=（aij）和b=（bij），通过重复建立矩阵乘积的计算（定义3.10之前），我们可以证明l（ab）=l（a）l（b）。

然后可以方便地使用矩阵表示法来描述线性映射l（a）的效果，如

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 |

引理6.4。让F:E×…×e→f为n-线性交替图。设（u1，…，un）和（v1，…，vn）为n向量的两个族，这样，

V1=A11U1+·····+AN1UN，

…Vn=a1nu1+·····+阿牛。

同等地，让

假设我们有

然后，

，

其中，和在所有置换πon 1，…，n范围内。

证据。用多线性展开f（v1，…，vn），得到形式的项和。

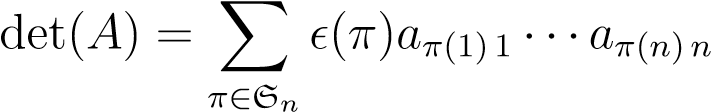
aπ（1）1···aπ（n）nF（uπ（1），…，uπ（n）），

对于所有可能的函数π：1，…，n→1，…，n。但是，因为f是交替的，所以只有π是置换的项是非零的。在命题6.1中，每个置换π都是转置数的乘积，而在命题6.2中，转置数的奇偶性只取决于π。然后，将命题6.3（3）应用于π中的每一个换位，我们得到

.

因此，我们得到了引理的表达式。

数量



实际上是a的行列式的值（我们稍后将看到，它也等于a>的行列式）。然而，直接使用上述定义是相当背道而驰的，我们将通过一条稍微间接的路线继续前进。

## 6.3行列式的定义

回想一下，在一个k域中，系数为n×n的全平方矩阵的集合用mn（k）表示。

定义6.4.行列式定义为任何映射

d:mn（k）→k，

当把它看作（kn）n上的映射时，也就是说，矩阵n列的映射，是n线性交替的，因此d（in）=1表示中的单位矩阵。等价地，我们可以考虑一个维数为n的向量空间e，一些固定基（e1，…，en），并定义

D:en→k

作为一个n-线性交替映射，使得d（e1，…，en）=1。

首先，我们将证明存在这样的映射D，使用一个归纳定义，也给出了一个递归方法来计算行列式。实际上，我们将定义一个族（d n）n≥1的（有限）映射集d:mn（k）→k。其次，我们将证明行列式实际上是唯一定义的，也就是说，我们将证明每个dn由一个映射组成。这将显示引理6.4的直接定义det（a）与归纳定义d（a）的等价性。最后，我们将用唯一性定理证明行列式的一些基本性质。

给定矩阵a∈mn（k），用a1，…，a n表示其n列。为了描述递归过程来定义一个行列式，我们需要一个小的概念。

定义6.5.对于n≥2的任意n×n矩阵，对于任意两个指数i，j（1≤i，j≤n），让aij为（n-1）×n-1）矩阵，该矩阵通过从a中删除第i行和第j列而获得，称为次矩阵：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误 | 网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 |
| 网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误 |
| 网络错误 | | | | |

定义6.6.对于每n≥1，我们将映射d:mn（k）→k的有限集dn归纳如下：

当n=1时，d1由单个映射d组成，因此，d（a）=a，其中a=（a），带有a∈k。

假设已经定义了dn−1，其中n≥2。然后，d n由所有的映射d组成，这样，对于某些i，1≤i≤n，

d（a）=（−1）i+1ai1d（a i 1）+······+（−1）i+naind（ain），

其中，对于每个j，1≤j≤n，d（aij）是将dn−1中的任何d应用于未成年人的结果。

Aij。

我们承认，对定义中的dn成员和dn−1成员使用相同的字母d可能会有点混淆。我们考虑使用下标来区分，但这似乎不必要地使事情复杂化。无论如何，我们不应该太担心，因为每个DN只包含一个映射。

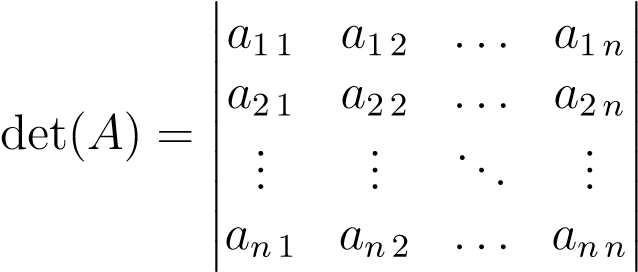
每个（−1）i+jd（aij）被称为aij的辅因子，d（a）的诱导表达式根据i行被称为d的拉普拉斯展开式。给定矩阵a∈mn（k），每个d（a）称为a的行列式。

我们可以把dn的每一个成员看作是一个计算a的“行列式”的算法。主要的一点是，这些算法使用所有可能的拉普拉斯行展开递归地计算一个行列式，都会得到相同的结果，det（a）。

我们将很快证明d（a）是唯一定义的（目前还不清楚dn由单个映射组成）。假设这个事实，给定一个n×n矩阵a=（aij），

|  |  |  |  |
| --- | --- | --- | --- |
| 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 |

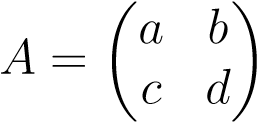
其行列式用d（a）或det（a）表示，或者更明确地用



首先，让我们先考虑一些例子。

例6.1。

1。当n=2时，如果

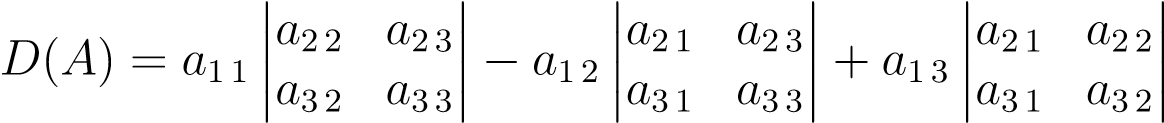


根据任何一行进行扩展，我们有

d（a）=ad−bc。

|  |  |  |
| --- | --- | --- |
| 网络错误 |  |  |
|  | 网络错误 |  |

根据第一排展开，我们有



也就是说，

d（a）=A11（A22A33−A32A23）−A12（A21A33−A31A23）+A13（A21A32−A31A22），给出了明确的公式

D（A）=A11A22A33+A21A32A13+A31A12A23−A11A32A23−A21A12A33−A31A22A13。

我们现在证明，每个d∈dn都是一个行列式（map）。

引理6.5。对于每个n≥1，对于定义6.6中定义的每个d∈dn，d是一个交替的多行映射，使得d（in）=1。

证据。通过对n的归纳，很明显d（in）=1。现在让我们证明D是多行的。让我们证明d在每一列中是线性的。考虑任何列k。因为

d（a）=（−1）i+1ai1d（a i 1）+··········+（−1）i+jaijd（aij）+·······（−1）i+naind（ain），

如果j=6k，那么通过归纳，d（aij）在k列是线性的，aij不属于k列，所以（−1）i+jaijd（aij）在k列是线性的。如果j=k，那么d（aij）不依赖于k=j列，因为aij是通过删除i行和j=k列而从a获得的，aij属于j=k列。因此，（）在K列是线性的。因此，在所有情况下，（-1）i+jaijd（aij）在K列是线性的，因此，d（a）在K列是线性的。

现在让我们证明d是交替的。假设A的两个相邻列相等，例如ak=ak+1。首先，设j=6 k，j=6 k+1。然后，矩阵aij有两个相同的相邻列，根据归纳假设，d（aij）=0。D（a）的其余条款为

（−1）i+kaikd（aik）+（−1）i+k+1aik+1d（aik+1）。

但是，这两个矩阵a i k和aik+1是相等的，因为我们假设a的k列和k+1是相同的，并且因为aik是通过删除行i和列k从a获得的，aik+1是通过删除行i和列k+1从a获得的。同样，aik=aik+1，因为a的k列和k+1相等。但是后来，

（−1）i+kaikd（aik）+（−1）i+k+1aik+1d（aik+1）=（−1）i+kaikd（aik）−（−1）i+kaikd（aik）=0.

这表明d是交替的，并完成了证明。

引理6.5表明行列式的存在。我们现在证明了他们的独特性。

定理6.6.对于每一个n≥1，对于每一个d∈dn，对于每一个矩阵a∈mn（k），我们有

，

其中，和在所有置换πon 1，…，n范围内。因此，dn由每n≥1的单个映射组成，该映射由上述显式公式给出。

证据。考虑kn的标准基（e1，…，en），其中（ei）i=1和（ei）j=0，对于j=6i，a的每列aj对应于矢量vj，其在基（e1，…，en）上的坐标是aj的分量，也就是说，我们可以写下

v1=a11e1+·····+an1en，

…Vn=a1ne1+·····+annen。

因为通过引理6.5，每个d是一个多行交替映射，通过应用引理6.4，我们得到

，

其中，和在所有置换πon 1，…，n范围内。但是d（e1，…，en）=d（in），根据引理6.5，我们得到d（in）=1。因此，

，

其中，和在所有置换πon 1，…，n范围内。

从现在开始，对于一个方阵的行列式，我们倾向于用符号det（a）代替d（a）。

注：行列式的几何解释非常有启发性。给定n个线性无关向量（u1，…，un），集合

pn=λ1u1+····+λnun 0≤λi≤1，1≤i≤n

称为Parallelotype。如果n=2，那么p2是一个平行四边形，如果n=3，那么p3是一个平行六面体，一个倾斜的盒子，它的三个角边是u1、u2、u3。然后，发现det（u1，…，un）是并行端口pn的有符号卷（其中volume表示n维卷）。本卷的符号说明PN在RN中的方向。

我们现在可以证明行列式的一些性质。

推论6.7。对于每个矩阵a∈mn（k），我们有det（a）=det（a>）。

证据。根据定理6.6，我们得到

，

其中，和在所有置换πon 1，…，n范围内。因为置换是可逆的，所以每个积

aπ（1）1···aπ（n）n

可以重写为

A1π−1（1）····ANπ−1（N）

既然）和的总和被1，…，n上的所有排列占据，我们有

，

其中π和σ在所有排列上都有范围。但立即证实

.

推论6.7的一个有用结果是矩阵的行列式也是其行的多行交替映射。这一事实，再加上矩阵的行列式是其列的多行交替映射这一事实，通常有助于在计算行列式时找到捷径。我们在下面的例子中说明这一点，这个例子在多项式插值中显示。

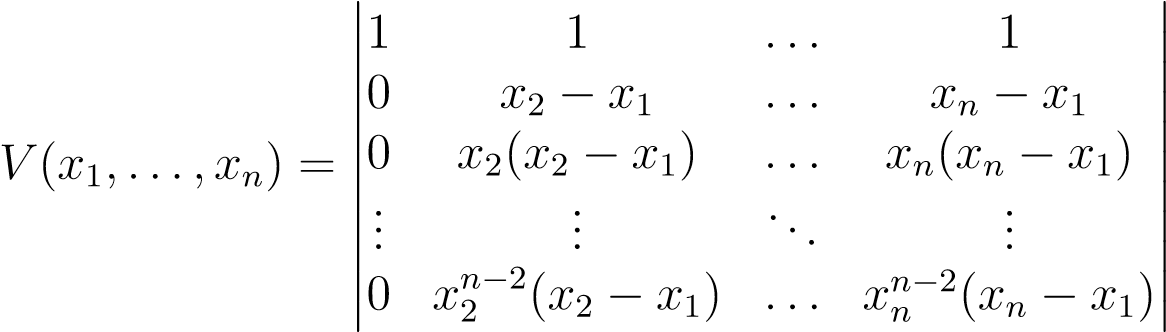
例6.2。考虑所谓的范德蒙行列式

.

我们声称

，

当n=1时，v（x1，…，xn）=1。我们通过n≥1上的归纳证明了这一点。案例n=1是显而易见的。假设n≥2。我们按照如下步骤进行：将第n-1行乘以x1，然后将其从第n行（最后一行）中减去，然后将第n-2行乘以x1，然后将其从第n-1行中减去，等等，将第i-1行乘以x1，然后将其从第i行中减去，直到我们到达第1行。我们得到下列行列式：



现在，根据第一列展开该行列式并使用多重线性，我们可以通过删除第一行和第一列获得的矩阵I i至1的列（Xi×X1），从而

V（x1，…，xn）=（x2−x1）（x3−x1）···（xn−x1）V（x2，…，xn）

这就建立了诱导步骤。

引理6.4可以很好地重新表述如下。

提案6.8.设f:e×…×e→f为n线性交替图。设（u1，…，un）和（v1，…，vn）为n向量的两个族，这样

V1=A11U1+·····+A1nun，

…Vn=An1U1+·····+阿牛。

|  |  |  |  |
| --- | --- | --- | --- |
| 网络错误 |  |  |  |
| 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误 | 网络错误  网络错误  网络错误 |

v1 U1

V2 U2

……=A…。VN联合国

那么，f（v1，…，vn）=det（a）f（u1，…，un）。

证据。与引理6.4的唯一区别是，这里我们使用的是>而不是a。因此，通过引理6.4和推论6.7，我们得到了期望的结果。

因此，我们得到了矩阵乘积的行列式是这些矩阵行列式的乘积这一非常有用的性质。

提案6.9.对于任意两个n×n矩阵a和b，我们有det（ab）=det（a）det（b）。

证据。我们使用命题6.8如下：让（e1，…，en）作为kn的标准基础，并让

w1 e1

w2 e2

。=ab…。

..wn en

然后，我们得到det（w1，…，wn）=det（ab）det（e1，…，en）=det（ab），

因为det（e1，…，en）=1。现在，让

v1 e1

V2 E2

…=B…，

γ

VN EN

我们得到

DET（v1，…，vn）=DET（b）

6.4。逆矩阵和行列式

从那以后

w1 v1

w2 v2

…=A…，

γ

WN VN

我们得到了det（w1，…，wn）=det（a）det（v1，…，vn）=det（a）det（b）。

应该注意的是，到目前为止，当k是一个交换环，而不一定是一个场时，这一部分的所有结果也成立。现在我们可以用行列式det（a）来描述一个n×n矩阵a是可逆的。

## 6.4逆矩阵和行列式

在接下来的两个部分中，k是一个交换环，当需要时，是一个场。

定义6.7.设k为交换环。给定一个矩阵a∈mn（k），让ae=（bij）是定义如下的矩阵：

bij=（−1）i+j det（aj i），第

aj i的辅因子。矩阵ae称为a的调整项，每个矩阵aj i称为矩阵a的次方。

注意指数的逆转

bij=（−1）i+j det（aj i）。

因此，ae是a元素的辅因子矩阵的转置。

我们有以下建议。

提案6.10.设k为交换环。对于每一个矩阵a∈mn（k），我们有

aae=aae=det（a）in.

因此，a是可逆的，如果det（a）是可逆的，那么a−1=（det（a））−1ae。

证据。如果ae=（bij）和aae=（cij），我们知道aae的第一行和第j列中的条目cij是

cij=ai1b1j+·····+aikbkj+····+ainbnj，

等于

ai1（−1）j+1 det（aj 1）+····+ain（−1）j+n det（aj n）。

如果j=i，那么我们根据第i行识别出det（a）的展开式：

cii=det（a）=a i 1（−1）i+1 det（ai1）+·····+a i n（−1）i+n det（ain）。

如果j=6i，我们可以用a的i行替换a的j行来形成矩阵a0。

现在，从A中删除行j和列k得到的矩阵aj k等于从a0中删除行j和列k得到的矩阵，因为a和a0只与第j行不同。因此，

，

我们有cij=ai1（−1）j+1 det（a0j 1）+······+ain（−1）j+n det（a0j n）。

然而，这是根据j行对det（a0）的展开，因为a0的j行等于a的i行，并且a0有两个相同的i和j行，因为det是行的交替映射（见前面的注释），所以我们得到det（a0）=0。因此，我们已经证明，当j=6i时，cii=det（a），cij=0，因此

.

从a的定义也可以明显看出，

.

然后，将参数的第一部分应用于>，我们将

A>AF>=DET（A>）输入，

从那以后，我们得到

，

也就是说，

（aae）>=det（a）in，

会产生

aae=Det（a）英寸，

从那以后。这证明了

aae=aae=det（a）in.

因此，如果DET（a）是可逆的，我们有一个−1=（DET（a））−1ae。相反，如果a是可逆的，从aa−1=in，根据命题6.9，我们得到了det（a）det（a−1）=1，det（a）是可逆的。

6.4。逆矩阵和行列式

当k是一个场时，元素a∈k是可逆的，当a=06。在这种情况下，命题的第二部分可以表述为a是可逆的iff det（a）=06。顺便注意一下，这种计算矩阵倒数的方法通常不实用。

我们现在考虑行列式在线性独立性和线性方程组求解中的一些应用。虽然这些结果适用于积分域上的矩阵，但它们的证明需要更复杂的方法（有必要使用积分域的分数域k）。因此，我们再次假设k是一个场。

设a为n×n矩阵，x a列为变量向量，b为另一列向量，a1，…，a表示a的列。观察方程ax=b的系统。

|  |  |  |  |
| --- | --- | --- | --- |
| 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 |

x1a1+····+xjaj+····+xnan=b，

因为在这两种情况下，第i行对应的方程都是

ai1x1+····+aijxj+····+ainxn=bi。

首先，我们用行列式来描述矩阵A的列向量的线性无关性。

提案6.11.给定一个n×n-矩阵a在一个k域上，a列的a1，…，a列与iff det（a）=det（a1，…，an）=0线性相关。相当地，a的秩n iff det（a）=06。

证据。首先，假设列a1，…，a的a是线性相关的。那么，有x1，…，xn∈k，这样

x1a1+····+xjaj+····+xnan=0，

其中xj=06，对于某些j。如果我们计算det（a1，…，x1a1+·················+xnan，…，an）=det（a1，…，0，…，an）=0，

其中0出现在j-th位置，通过多重性，所有包含两个相同列k=6j的项都消失了，我们得到

xj det（a1，…，an）=0.

因为xj=06，k是一个字段，所以我们必须有det（a1，…，an）=0。

相反，我们表明，如果列a1，…，a的a是线性无关的，那么det（a1，…，an）=06。如果a1，…，a的a列是线性无关的，那么它们构成了kn的基，我们可以用下式表示kn的标准基（e1，…，en）。

|  |  |  |  |
| --- | --- | --- | --- |
| 网络错误 |  |  |  |
| 网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 |

对于一些矩阵b=（bij），通过命题6.8，我们得到

DET（e1，…，en）=DET（b）DET（a1，…，an）

由于det（e1，…，en）=1，这意味着det（a1，…，an）=0（6，det（b）=0）6。对于第二个断言，回想一下，矩阵的秩等于线性无关列的最大数目，并且结论是清楚的。

如果我们把6.11和10.14结合起来，我们就得到了矩阵的秩的下列标准。

提案6.12。给定一个K域上的任意m×n矩阵a（通常k=r或k=c），a的秩是最大自然数r，这样通过选择a的r行和r列得到a的r×r子矩阵b，这样det（b）=06。

六

## 6.5线性方程组和行列式

我们现在描述一个形式为ax=b的线性方程组有唯一解的时候。

提案6.13。给定一个域k上的n×n矩阵a，以下属性成立：

1. 对于每个列向量b，都有一个唯一的列向量x，这样ax=b iff，ax=0的唯一解是平凡向量x=0，iff det（a）=06。
2. 如果det（a）=06，则由表达式给出ax=b的唯一解。

，

被称为克莱默法则。

1. 线性方程组ax=0有一个非零解iff det（a）=0。

6.6。线性映射的行列式

证据。假设ax=b有一个解X0，假设y=0，y=0.6，那么，

A（X0+Y）=Ax0+Ay=Ax0+0=B，

X0+Y=6X0是ax=b的另一个解，包含了ax=b有一个解X0的假设。因此，ax=0只有平凡解。现在，假设ax=0只有平凡的解。这意味着a1，…，a列是线性独立的，根据命题6.11，我们得到了det（a）=0.6。最后，如果det（a）=06，根据命题6.10，这意味着a是可逆的，然后，对于每个b，a x=b等于x=a−1b，这表明ax=b有一个单独的解。

1. 假设ax=b。如果我们计算det（a1，…，x1a1+·····+xjaj+·····+xnan，…，an）=det（a1，…，b，…，an），

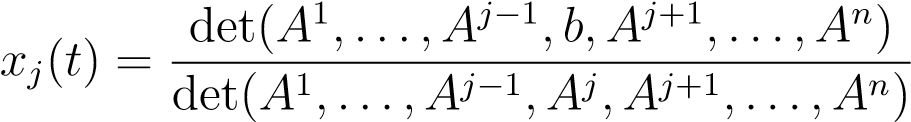
当b出现在j-th位置时，通过多线性，所有包含两个相同列k=6j的项都消失了，我们得到

xj det（a1，…，an）=det（a1，…，aj−1，b，aj+1，…，an）

对于每个j，1≤j≤n。由于我们假设det（a）=det（a1，…，an）=06，我们得到了所需的表达式。

1. 注意，ax=0有一个非零解iff a1，…，an是线性相关的（如在命题6.11的证明中所观察到的），通过命题6.11，它等于det（a）=0。

尽管克莱默法则令人满意，但用上述表达式求解线性方程组通常是不切实际的。然而，这些公式暗示了一个有趣的事实，即系统ax=b的解在a和b中是连续的。如果我们假设a中的项是连续函数aij（t），b中的项也是实参数t的连续函数bj（t），因为行列式是其项的多项式函数，表达式



是多项式的比值，因此只要det（a（t））不为零，它也是连续的。同样，如果函数aij（t）和bj（t）是可微的，那么xj（t）也是可微的。

## 6.6线性图的行列式

我们用线性映射f:e→e的行列式的概念来结束这一章。

给定一个有限维n的向量空间e，给定e的基（u1，…，un），对于每一个线性映射f:e→e，如果m（f）是f w.r.t的矩阵，基（u1，…，un），我们可以定义det（f）=det（m（f））。如果（v1，…，vn）是e的任何其他基，如果p是基矩阵的变化，根据推论4.5，f关于基（v1，…，vn）的矩阵是p−1m（f）p。现在，根据命题6.9，我们得到了det（p−1m（f）p）=det（p−1）det（m（f））det（p）=det（p−1）det（p）det（m（f））=det（m（f））。

因此，det（f）确实独立于e的基础。

定义6.8.给定一个有限维的向量空间e，对于任意线性映射f:e→e，我们将f的行列式det（f）定义为f的矩阵在任何基础上的行列式det（m（f））（因为在定义之前的讨论中，该行列式不依赖于基）。

那么，我们有以下的建议。

提案6.14.对于任意有限维n的向量空间e，线性映射f:e→e是可逆的iff det（f）=06。

证据。线性映射f:e→e是可逆的，如果它的矩阵m（f）在任何基上都是可逆的（通过命题4.2），iff det（m（f））=06，通过命题6.10。

给定一个有限维N的向量空间，可以很容易地看出，双射线性映射的集合f:e→e使得det（f）=1是一个组合下的群。这个群是一般线性群gl（e）的一个子群。它被称为特殊线性群（e），用sl（e）表示，当e=k n时，用sl（n，k）表示，甚至用sl（n）表示。

## 6.7凯莱-汉密尔顿定理

我们以一个有趣而重要的6.10命题的应用，即凯莱-汉密尔顿定理，来结束这一章。本节的结果适用于任何交换环k上的矩阵。首先，我们需要矩阵特征多项式的概念。

定义6.9.如果k是任意交换环，对于每一个n×n矩阵a∈mn（k），a的特征多项式pa（x）是行列式。

Pa（x）＝DET（Xi—a）。

特征多项式pa（x）是k[x]中的一个多项式，不定x中的多项式环的系数是k。例如，当n=2时，如果

然后

6.7。凯莱-汉密尔顿定理

我们可以用矩阵a代替多项式pa（x）中的变量x，得到矩阵pa。

pa（x）=xn+c1xn−1+····+cn，

然后

PA=AN+C1AN−1+·····+CNI。

我们有以下显著的定理。

定理6.15。（Cayley–Hamilton）如果k是任何交换环，对于每个n×n矩阵

a∈mn（k），如果我们让

pa（x）=xn+c1xn−1+····+cn

是的特征多项式，然后

pa=an+c1an−1+····+cni=0。

证据。我们可以把矩阵B＝Xi-A作为矩阵的系数在多项式环k[x]中看，然后我们可以形成矩阵BE，它是元素B的辅因子矩阵的转置。Te B作为

be=xn−1b0+xn−2b1+···+bn−1，

对于一些系数为k的矩阵b0，…，bn-1。例如，当n=2时，我们有

.

根据6.10号提案，我们

bbe=Det（b）i=PA（x）i。

另一方面，我们有

BBE＝（Xi～A）（Xn＝1B0+Xn＝2B1+···+Xn·j＝1BJ+···+BN×1）；

把右手边乘以，我们得到

bbe=xnd0+xn−1d1+···+xn−jdj+····+dn，

具有

d0=b0 d1=b1−ab0

…

dj=bj−abj−1

…

dn−1=bn−1−abn−2 dn−abn−1。

自从

pa（x）i=（xn+c1xn−1+····+cn）i，

平等

xnd0+xn−1d1+····+dn=（xn+c1xn−1+···+cn）i

是两个矩阵之间的等式，因此它要求所有对应的项都相等，并且由于这些项是多项式，因此这些多项式的系数必须相同，这相当于一组方程

i=b0 c1i=b1−ab0

…

Cji=bj−abj−1

…

cn-1i=bn-1−abn-2

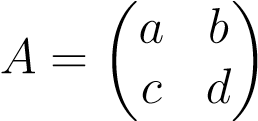
CNI=−ABN−1，

对于所有j，1≤j≤n-1。如果我们将第一个方程乘以a，最后一个方程乘以i，通常（j+1）th乘以−j，当我们把所有这些新方程相加时，我们看到右边加起来是0，我们得到了我们想要的方程。

AN+C1AN−1+·····+CNI=0，

如要求。

作为一个具体的例子，当n=2时，矩阵



满足方程

A2−（A+D）A+（AD−BC）i=0.

大多数读者可能会发现定理6.15的证明相当聪明，但非常神秘，毫无动力。概念上的困难在于，我们真的需要了解一个变量中的多项式是如何“作用”于向量的，就矩阵A而言。这可以做到，并产生一个更“自然”的证明。实际上，如果我们把自己从矩阵中解放出来，而不是考虑有限维向量空间e和一些给定的线性映射f:e→e，那么推理更简单，更一般。给定任意多项式p（x）=a0xn+a1xn−1+········+an和k域中的系数，我们定义线性映射p（f）：e→e by

p（f）=a0fn+a1fn−1+····+anid，

6.7。凯莱-汉密尔顿定理

式中，f k=f·····f，f与其自身的k-折叠组成。注意

p（f）（u）=a0fn（u）+a1fn−1（u）+····+anu，

对于每一个向量u∈e，我们用多项式定义了一种新的标量乘法·：k[x]×e→e，对于每一个多项式p（x）∈k[x]，对于每一个u∈e，

p（x）·u=p（f）（u）。

很容易证实这是一个“良好的行动”，也就是说

P·（U+V）=P·U+P·V（P+Q）·U=P·U+Q·U

（PQ）·U=P·（Q·U）1·U=U，

对于所有p，q∈k[x]和所有u，v∈e，用这个新的标量乘法，e是k[x]模。如果p=λ只是k中的一个标量（阶数为0的多项式），那么

λ·u=（λid）（u）=λu，

也就是说，K通过标量乘作用于E。如果p（x）=x（单项x），那么

x·u=f（u）。

现在，如果我们选取一个基（e1，…，en），如果一个多项式p（x）∈k[x]具有

p（x）·ei=0，i=1，…，n，

这意味着p（f）（e i）=0，对于i=1，…，n，这意味着线性映射p（f）在e上消失。我们也可以像在6.2节中所做的那样，检查如果a和b是两个n×n矩阵，如果（u1，…，un）是任何n个向量，那么

.

这就提出了我们第二次证明凯莱-汉密尔顿定理的攻击计划。为了简单起见，我们证明了场上向量空间的定理。证明了交换环上自由模的存在。

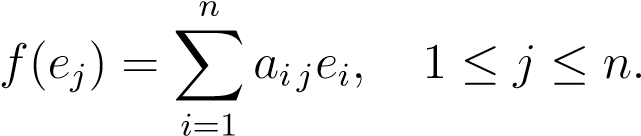
定理6.16。（Cayley–Hamilton）对于场k上的每个有限维向量空间，对于每个线性映射f:e→e，对于每个基（e1，…，en），如果a是基（e1，…，en）上f的矩阵，如果

pa（x）=xn+c1xn−1+····+cn

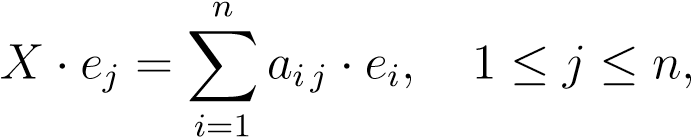
是的特征多项式，那么

pa（f）=fn+c1fn−1+····+cnid=0。

证据。由于a的列由表示在基（e1，…，en）上的向量f（ej）组成，我们有



利用k[x]对e的作用，上述方程可以表示为



会产生

，1≤j≤n。

观察特征多项式的转置出现，因此上述系统可以写成

.

如果我们假设B= Xi→a>，那么如前一个证明，如果是B的余因子矩阵的转置，则

BBE＝DET（b）i＝DET（Xi→a>）i＝DET（Xi→a）i＝PAI。

但从那时起

，

因为Be是一个矩阵，其项是k[x]中的多项式，所以在左边乘Be是有意义的，我们得到

；

也就是说，

pa·ej=0，j=1，…，n，

如权利要求所述，证明pa（f）=0。

6.8。永久物

如果k是一个场，那么线性映射f:e→e的特征多项式与e中选择的基（e1，…，en）无关。为了证明这一点，观察到对于某些不可数矩阵p，f在另一个基上的矩阵将是p-1ap的形式，然后

属性（XI）

= DET（p＝1（Xi（a））p）＝DET（p＝1）DET（Xi（a））DET（p）＝DET（Xiα）。

因此，线性映射的特征多项式是f的固有性质，用pf表示。

线性映射F的特征多项式的零（根）称为F的特征值，在理论和应用中起着重要作用。稍后我们将回到这个主题。

## 6.8永久性

回想一下，n×n矩阵行列式的显式公式是

.

如果我们从上面的公式中去掉每个排列的符号），我们得到一个称为永久数的量：

每（a）=x aπ（1）1···aπ（n）n。

π∈sn

柯西早在1812年就研究了永久性和决定因素。从上述定义可以清楚地看出，永久形式是多行对称形式。我们也有

每（a）=每（a>），

以及拉普拉斯展开式的以下无符号版本：

每（a）=ai1 per（ai1）+····+aijper（aij）+·····+aiper（ain）

然而，对于i=1，…，n.与行列式不同，行列式有明确的几何解释，作为有符号的体积，永久数没有任何自然的几何解释。此外，行列式可以有效地评估，例如使用行缩减梯队形式的转换，但是计算永久行列式是困难的。

永久的结果是有各种组合的解释。其中之一就是我们现在讨论的二部图的完全匹配。

回想一个二部（无向）图g=（v，e）是一个图，它的节点集v可以分成两个非空的不相交的子集v1和v2，这样每个边e∈e在v1中有一个端点，在v2中有一个端点。图6.8显示了一个具有14个节点的双平面图的示例；其节点被划分为两组x1、x2、x3、x4、x5、x6、x7和y1、y2、y3、y4、y5、y6、y7。

*x*

1

*x*

2

*x*

3

*x*

4

*x*

5

*x*

6

*x*

7

*y*

1

*y*

2

*y*

3

*y*

4

*y*

5

*y*

6

*y*

7

图6.1：二部图G。

图G=（V，E）（是否二部）中的匹配是一组成对的非相邻边，这意味着M中没有两个边共享一个公共顶点。完美匹配是这样一种匹配，即v中的每个节点都与匹配m中的某个边相关联（v中的每个节点都是m中某个边的端点）。图6.8显示了二部图G中的完全匹配（红色）。

*x*

1

*x*

2

*x*

3

*x*

4

*x*

5

*x*

6

*x*

7

*y*

1

*y*

2

*y*

3

*y*

4

*y*

5

*y*

6

*y*

7

图6.2：二部图G中的完美匹配。

显然，只有当二部图的一组节点在两个大小相等的块（如x1，…，xm和y1，…，ym）中有一个分区时，才能存在完全匹配。然后，在完美匹配和双射π之间有一个双射：{x1，…，xm }，{ y1，…，ym }，使得π（Xi）＝yj-IFF，在Xi和YJ之间有一个边。

现在，每一个二部图G，其节点被划分成两个大小相等的集合，如上图所示，用m×m矩阵a=（aij）表示，这样aij=1iff有一条边

6.9。进一步阅读

在XI和YJ之间，而AIJ＝0。利用双射π：x1，…，xm→y1，…，ym的完美机械的解释，我们看到表示二部图g的（0,1）矩阵a的永久per（a）计数g中的完美匹配数。

在1979年出版的一篇著名论文中，LeslieValiant证明了计算永久性是一个P-完全问题。这些问题被怀疑是难以解决的。众所周知，如果存在一个多项式时间算法来解决一个p-完全问题，那么我们将得到p=np，这被认为是非常不可能的。

另一种对恒量的组合解释可以用不同代表的系统来解释。给定一个有限集s，让（a1，…，an）是s的任何非空子集序列（不一定是不同的）。集合a1，…，a n是n个不同元素（a1，…，an）的序列，a i∈ai表示i=1，…，n，集合序列的sdr个数在组合数学中起着重要作用。现在，如果s=1,2，…，n并且如果我们关联到s的非空子集的任何序列（a1，…，an），那么矩阵a=（ai j）定义为aij=1，如果j∈ai和aij=0，那么永久per（a）计算集合a1，…，an的sdr的数量。

用SDR解释永久性可以用来证明各种矩阵的永久性的界限。感兴趣的读者可参考van Lint和Wilson[174]（第11章和第12章）。特别是，在第12章中给出了一个被称为范德瓦尔登猜想的定理的证明。这个定理指出，对于任何n×n矩阵a，其中所有的行和和列和都是1（双随机矩阵），我们有

每，（

其中所有项都等于1/n的矩阵的等式。

## 6.9进一步读数

第3-5章和第6章所述材料的详细论述可在Strang[165，164]、Lax[110]、Lang[106]、Artin[7]、Mac Lane和Birkhoff[115]、Hoffman和Kunze[99]、Bourbaki[25，26]、van der Waerden[173]、Serre[151]、Horn和Johnson[92]和Bertin[15]中找到。线性代数的这些概念很好地应用于古典几何，见Berger[11，12]、Tisseron[170]和Dieudonn'e[50]。

162第6章。行列式

第七章

# 高斯消去、Lu因子分解、Cholsky因子分解、缩减行梯队形式

在这一章中，我们假设所有的向量空间都在r域上，所有的结果不依赖于r的次序，也不依赖于任意域的平方根。

## 7.1激励示例：曲线插值

曲线插值是计算机图形学和机器人学（路径规划）中经常出现的问题。有很多方法可以解决这个问题，在本节中，我们将描述一个使用三次样条曲线的解决方案。这种样条由三次B’Ezier曲线组成。它们经常被使用，因为它们的实现成本较低，并且比二次B’ezier曲线具有更大的灵活性。

三次B’ezier曲线c（t）（在r2或r3中）由四个控制点列表指定。

（b0，b2，b2，b3），并由方程式参数化给出

c（t）=（1 t）3 b0+3（1 t）2tb1+3（1 t）t2 b2+t3 b3.

显然，t∈[0,1]，点c（t）属于控制点b0、b1、b2、b3的凸壳。多项式

（1−t）3、3（1−t）2t、3（1−t）t2、t3

是三阶伯恩斯坦多项式。

通常，我们只对区间内t值对应的曲线段感兴趣[0,1]。尽管如此，控制点的位置仍然会严重影响曲线段的形状，甚至可能有一个自相交；请参见图7.1、7.2、7.3，以说明各种配置。

一百六十三

*b*

0

*b*

1

*b*

2

*b*

3

图7.1：“标准”b’ezier曲线。

*b*

0

*b*

1

*b*

2

*b*

3

图7.2：带有拐点的B’ezier曲线。

7.1。激励示例：曲线插值

*b*

0

*b*

1

*b*

2

*b*

3

图7.3：自交b’ezier曲线。

插值问题需要找到通过某些给定数据点的曲线，并可能满足一些额外的约束。

B’ezier样条曲线f是由B’ezier曲线（如c1，…，cm（m≥2））的曲线段组成的曲线。我们假设f在[0，m]上定义，因此对于i=1，…，m，

f（t）=Ci（t−i+1），i−1≤t≤i。

通常，在任意两个连接点之间，即在任意两点Ci（1）和Ci+1（0）之间，对于i=1，…，m-1，需要一定的平滑度。我们要求ci（1）=ci+1（0）（c0连续性），通常情况下，ci在1和ci+1在0的导数等于二阶导数。这被称为C2连续性，它确保切线和曲率一致。

有许多插值问题，我们认为其中一个最常见的问题可以表述为：

问题：给定n＋1个数据点x0，…，xn，找到一个C2三次样条曲线F，使得f i（i）＝Xi为i，0±i±n（n＝2）。

解决这个问题的一种方法是找到n+3辅助点d−1，…，dn+1，称为de boor控制点，从中可以找到n b´ezier曲线。事实上，

D−1=X0和DN+1=XN

所以我们只需要找到n+1点，d0，…，dn。

结果表明，N-B’ezier曲线上的C2连续性约束只产生N-1方程，因此可以任意选择D0和DN。在实践中，根据不同的端部条件，如规定的X0和XN速度，选择D0和DN。目前，我们假定给出了d0和dn。

图7.4说明了一个涉及n+1=7+1=8个数据点的插值问题。任意选择控制点D0和D7。

*x*

0

=

*d*

−

1

*x*

1

*x*

2

*x*

3

*x*

4

*x*

5

*x*

6

*x*

7

=

*d*

8

*d*

0

*d*

1

*d*

2

*d*

3

*d*

4

*d*

5

*d*

6

*d*

7

图7.4：通过点X0、X1、X2、X3、X4、X5、X6、X7的C2三次插值样条曲线。

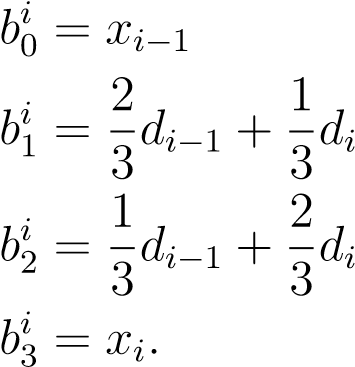
可以看出，d1，…，dn-1由线性系统给出。

.

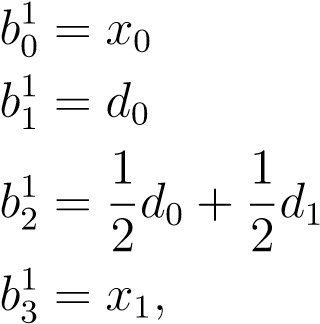
稍后我们将证明，上面的矩阵是可逆的，因为它是严格对角占优的。

一旦上述系统被解决，b’ezier cubics c1，…，cn的确定如下：

（我们假设n≥2）：对于2≤i≤n−1，控制点（由



控制点（由



控制点（由

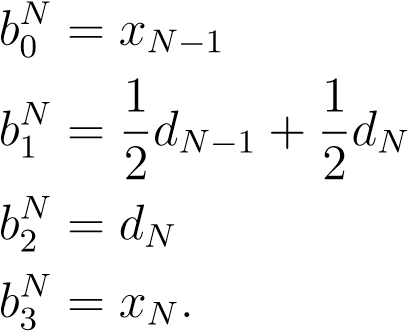


图7.5说明了n=7的过程样条插值。

我们现在将描述求解线性系统的各种方法。由于上述系统的矩阵是三对角的，所以有专门的方法比一般的方法更有效。我们将讨论其中的一些方法。

## 7.2高斯消元

设a为n×n矩阵，设b∈rn为n维向量，并假定a是可逆的。我们的目标是解决系统a x=b。由于假设a是可逆的，我们知道这个系统有一个唯一的解决方案x=a-1b。经验表明，揭示了两个反直觉的事实：

1. 应避免显式计算a的逆a−1。这是无效的，因为它等于求解n个线性系统au（j）=ej，对于j=1，…，n，其中ej=（0，…，1，…，0）是rn的第j个规范基向量（其中1是第j个槽）。通过这样做，我们可以用n个系统的分辨率来代替单个系统的分辨率，我们仍然需要用a−1乘以b。

图7.5:x0、x1、x2、x3、x4、x5、x6、x7的c2三次插值，带有相关的颜色编码b'ezier cubics。

1. 我们不能通过计算行列式（使用克莱默公式）来求解（大型）线性系统，因为这种方法需要大量的加法（resp）。乘法）与（n+1）成比例！（响应（N+2）！）.

大多数直接方法（与迭代方法相反，寻找解的近似值）基于的关键思想是，如果a是上三角矩阵，这意味着对于1≤j<i≤n（resp），aij=0。下三角，这意味着对于1≤i<j≤n，aij=0，那么计算解x是微不足道的。实际上，假设A是一个上三角矩阵

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误 |

那么det（a）=a11a22·······a n n=06，这意味着i=1，…，n时aii=06，我们可以通过反替换从下到上求解系统ax=b。也就是说，首先我们从最后一个方程计算xn，然后将xn的这个值插入到最后一个方程的下一个方程中，并从中计算xn−1，依此类推。这将产生

|  |  |  |
| --- | --- | --- |
| 网络错误 | 网络错误 |  |
| 网络错误 | 网络错误  网络错误 |  |
| 网络错误 | 网络错误 | 网络错误 |

注意，可以避免使用行列式来证明，如果a是可逆的，那么a i i=06对于i=1，…，n。事实上，可以直接（通过归纳）证明，当上（或下）三角形矩阵的所有对角项都不为零时，上（或下）三角形矩阵是可逆的。

如果A是下三角，我们用正替换法从上到下求解系统。

因此，我们需要的是一种将矩阵转换为上三角形式的等价矩阵的方法。这可以通过消除来实现。让我们用下面的例子来说明这个方法：

2x+y+z=5

4x−6y=−2

-2X+7Y+2Z=9。

我们可以从第二个和第三个方程中去掉变量x，如下所示：从第二个方程中减去第一个方程的两倍，然后将第一个方程添加到第三个方程中。我们得到了新的系统

2x+y+z=5

−8Y−2Z=−12 8Y+3Z=14。

这一次，我们可以通过将第二个方程添加到第三个方程来消除第三个方程中的变量y：

2x+y+z=5

−8Y−2Z=−12 Z=2.

最后一个系统是上三角形。使用反替换，我们找到了解决方案：z=2，y=1，x=1。

请注意，我们只执行了行操作。一般的方法是使用简单的行操作（即在矩阵的另一行中添加或减去一行的倍数）迭代消除变量，同时将这些操作应用于向量b，以获得一个系统max=mb，其中ma是上三角形。这种方法叫做高斯消元法。但是，该方法在所有情况下都需要一个额外的扭曲：可能需要排列行，如下面的示例所示：

x+y+z=1 x+y+3z=1

2x+5y+8z=1.

为了从第二行和第三行中减去X，我们从第二行中减去第一行，从第三行中减去第一行的两倍：

X+Y+Z=1

2Z=0 3Y+6Z=−1.

现在的问题是y不会出现在第二行；所以，我们不能通过在第三行中加上或减去第二行的倍数来消除y。补救办法很简单：排第二排和第三排！我们得到系统：

X+Y+Z=1

3Y+6Z=−1 2Z=0，

已经是三角形了。另一个需要排列的例子是：

Z=1−2X+7Y+2Z=1 4X−6Y=−1。

首先我们排列第一排和第二排，得到

-2X+7Y+2Z=1 Z=1

4x−6y=−1，

然后我们将第一行的两倍添加到第三行，得到：

-2X+7Y+2Z=1 Z=1

8Y+4Z=1.

我们再次排列第二排和第三排，得到

-2X+7Y+2Z=1

8Y+4Z=1 Z=1，

上三角形系统。当然，在这个例子中，z已经被解决了，我们可以先消除它，但是对于一般的方法，我们需要以一种系统的方式进行。

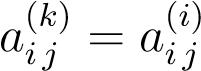
我们现在描述了应用于线性系统ax=b的高斯消元方法，其中a被假定为可逆的。我们使用变量k来跟踪消除的阶段。最初，k=1。

1. 第一步是在a的第一列中选择一些非零项ai1。这样的项必须存在，因为a是可逆的（否则，a的第一列将是零向量，a的列将不是线性独立的）。同样地，我们会得到det（a）=0。这种单元的实际选择对该方法的数值稳定性有一定的影响，但稍后将对此进行研究。目前，我们假设做出了一些任意的选择。所选元素被称为消除步骤的支点，并表示为π1（因此，在第一步中，π1=ai1）。
2. 接下来，我们将与第一行对应的轴排列行（i）。这样的一步叫做旋转。所以在这种排列之后，第一行的第一个元素是非零的。
3. 现在，我们通过向这些行添加第一行的适当倍数，从除第一行之外的所有行中消除变量x1。更精确地说，我们将−ai1/π1乘以第一行，得到i=2，…，n的第i行。在这个步骤的末尾，第一列中的所有条目都是零，除了第一列。
4. 将k增加1。如果k=n，停止。否则，k<n，然后在（n−k+1）×（n−k+1）子系统上重复步骤（1）、（2）、（3），通过从当前系统中删除第一个k−1行和k−1列获得。

如果我们让a1=a和）是K−1消除步骤（2≤k≤n）后得到的矩阵，则将kth消除步骤应用于形式的矩阵ak。

.

事实上，请注意



对于所有i，j，其中1≤i≤k−2和i≤j≤n，因为在（k−1）第（k−1）步之后，第一个k−1行保持不变。

稍后我们将证明det（ak）=？det（a）。因此，AK是可逆的。AK是可逆的事实，如果A是可逆的，也可以不用行列式来表示，因为存在一些可逆矩阵mk，这样AK=mk a，我们很快就会看到。

由于ak是可逆的，所以k≤i≤n的某些项是非零的。否则，ak的前k列中的最后n−k+1项将为零，而ak的前k列将在rk−1中生成k向量。但AK的前k列是线性相关的，AK不可逆，这是一个矛盾。这种情况由以下n=5和k=3的矩阵说明：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误 | 网络错误  网络错误  网络错误  网络错误 |

上述矩阵的前三列是线性相关的。

因此可以选择其中一个k≤i≤n的项作为轴，并将第k行与第i行进行排序，得到矩阵）。新的数据透视是，我们将K列中的条目i=k+1，…，n加上时间行k到行i，归零。在这个步骤的最后，我们得到了ak+1。观察AK的前k-1行与AK+1的前k-1行相同。

高斯消元过程如下图所示：

×××××××。

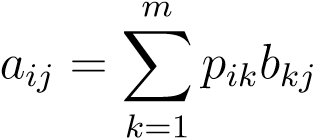
×××=0×××=0×。

 

×0×0×0×0×0×0×0 0×0

## 7.3基本矩阵和行操作

很容易知道什么样的矩阵执行高斯消去过程中使用的基本行操作。关键是，如果a=p b，其中a，b是m×n矩阵，p是维m的平方矩阵，如果（通常）我们用a1，…，am和b1，…，bm表示a和b的行，那么公式



给出a中的（i，j）th项表明a的第i行是b行的线性组合：

ai=pi1b1+·····+pimbm。

因此，左边的矩阵乘以一个正方形矩阵就可以执行行操作。类似地，右边的矩阵乘以一个正方形矩阵，就可以执行列操作。

第k行与第i行的置换是通过将左边的a乘以换位矩阵p（i，k）来实现的，换位矩阵p（i，k）是从单位矩阵中获得的矩阵。

7.3。初等矩阵与行运算

通过排列行i和k，即，

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误 | 网络错误  网络错误 | 网络错误 | 网络错误 | 网络错误 | 网络错误  网络错误 | 网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 |

例如，如果m=3，

，

然后

.

观察Det（P（I，K））=1。此外，p（i，k）是对称的（p（i，k）>=p（i，k）），p（i，k）−1=p（i，k）。

在置换步骤（2）中，如果需要对第k行和第i行进行置换，则矩阵a在左边乘以矩阵pk，使pk=p（i，k），否则我们设置pk=i。

将β乘以第j行到第i行（i=6 j），将左边的a乘以初等矩阵，

ei，j；β=i+βeij，

哪里

= J

，即，

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误 | 网络错误  网络错误 | 网络错误 | 网络错误 | 网络错误 | 网络错误 | 网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误 | 网络错误 | 网络错误 | 网络错误 | 网络错误 | 网络错误 | 网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 |

在左边，i>j，在右边，i<j。索引i是被乘法改变的行的索引。例如，如果m=3，我们想将第1行添加到第3行，因为β=2，j=1，i=3，我们形成

，

然后计算

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 网络错误  网络错误  网络错误  网络错误  网络错误 | 1. 网络错误 2. 网络错误   网络错误 | 网络错误 | 网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 |

2B11+B31 2B12+B32········2B1N+B3N

观察ei，j的倒数；β=i+βeij为ei，j；−β=i−βeij，且det（ei，j；β）=1。因此，在步骤3（消除步骤）中，矩阵A左乘形式为ei，k；βi，k，i>k的矩阵的乘积ek。

因此，我们看到

AK+1=EKPAK，

然后

AK=EK−1PK−1 E1P1A。

这证明了之前的声明，对于某些可逆矩阵mk，我们可以选择

mk=ek−1pk−1···e1p1，

可逆矩阵的乘积。

DET（p（i，k））=1和DET（ei，j；β）=1这一事实立即意味着上述事实：我们一直

Det（ak）=？Det（a）。

此外，因为

AK=EK−1PK−1···E1P1A

由于高斯消去停止k=n，矩阵

an=en−1pn−1····e2p2e1p1a

是上三角形。还要注意，如果m=en−1pn−1····e2p2e1p1，那么det（m）=±1，det（a）=±det（an）。

矩阵p（i，k）和ei，j；β称为初等矩阵。我们可以将上述讨论概括为以下定理：

定理7.1.（高斯消去）设a为n×n矩阵（可逆与否）。然后有一些可逆矩阵m，所以u=ma是上三角形。支点都是非零的，如果a是可逆的。

证据。我们已经证明了当a可逆时的定理，以及最后一个断言。现在a是奇异的，如果某个支点是零，比如在消去的阶段k。如果是的话，我们必须有；但是在这种情况下，ak+1=ak，我们可以选择pk=ek=i。

注：显然，矩阵m可以计算为

m=en−1pn−1···e2p2e1p1，

但这个表达是没有用的。事实上，我们需要的是m−1；当不需要排列时，结果表明m−1可以立即从矩阵Ek的逆矩阵中获得，并且不需要乘法。

注：不需要求可逆矩阵m，使m a为上三角形，我们可以求可逆矩阵m，使ma为对角矩阵。只需要简单地改变高斯消元。在每个阶段k，在找到轴并执行轴转动后，如有必要，除了将第k行的适当倍数添加到第k行以下的行中，以便使i=k+1，…，n的k列中的条目归零，还将第k行的适当倍数添加到第k行以上的行中。k为使i=1，…，k-1列中的条目归零。除i<k外，这些步骤还通过将左边的矩阵乘以初等矩阵ei，k；βi，k来实现，因此这些矩阵不是下三角矩阵。然而，在这个过程的最后，我们发现an=ma是一个对角矩阵。

这种方法称为高斯-乔丹因式分解。由于该方法比高斯消元法更昂贵，因此在实际应用中应用不多。然而，高斯-乔丹因式分解可用于计算矩阵A的逆矩阵。实际上，我们通过解系统ax（j）=ej（其中ej是rn的第j个标准基向量）找到了a−1的第j列。通过应用高斯-乔丹，我们得到了dj x（j）=mjej形式的系统，其中dj是一个对角矩阵，我们可以立即计算x（j）。

它还需要讨论支点的选择，以及保证在高斯消元过程中不需要排列的条件。我们首先说明了可逆矩阵具有Lu因子分解的必要和充分条件（即，高斯消元不需要旋转）。

## 7.4 LU因子分解

定义7.1.我们假设一个可逆矩阵A有一个Lu因式分解，如果它可以写成a=Lu，其中u是上三角可逆的，l是下三角的，其中l i i=1代表i=1，…，n。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 网络错误 |  |  |  |  |
| 网络错误 | 网络错误  网络错误  网络错误  网络错误 | 网络错误 | 网络错误 | 网络错误  网络错误 |
|  |  | 网络错误 | 网络错误 | 网络错误 |
|  | |

对角项等于1的下三角矩阵称为单位下三角矩阵。给定一个n×n矩阵a=（a i j），对于1≤k≤n的任何k，让a（1:k，1:k）表示其项为aij的a的次矩阵，其中1≤i，j≤k。例如，如果a是5×5

提案7.2.设a为可逆n×n矩阵。那么A有一个Lu分解

a=lu如果每个矩阵a（1:k，1:k）对于k=1，…，n是可逆的。此外，当a有lu因子分解时，我们有

Det（A（1:K，1:K））=π1π2··································

其中πk是K-1消除步骤后获得的轴。因此，kth轴由

γ

如果k=1，则a11=det（a（1:1,1:1））。

如果k=2，…，n.

证据。首先假设a=lu是a的lu因子分解。

，

其中，l1、l4为单位下三角，u1、u4为上三角。（注：a（1:k，1:k）、l1和u1是k×k矩阵；a2和u2是k×（n-k）矩阵；a3和l3是（n-k）×k矩阵；a4、l4和u4是（n-k）×（n-k）矩阵。）

A（1:k，1:k）=l1u1，

由于u是可逆的，u1也是可逆的（u的行列式是u中对角线项的乘积，是u1和u4中对角线项的乘积）。因为l1是可逆的（因为它的对角线项等于1），我们看到a（1:k，1:k）对于k=1，…，n是可逆的。

相反，假设a（1:k，1:k）对于k=1，…，n是可逆的。我们只需要证明高斯消元不需要旋转。通过K上的归纳证明了第k步不需要转动。

因为a（1:1,1:1）=（a11），所以a11=06，所以k=1。假设前k-1步（2≤k≤n-1）不需要旋转。在这种情况下，我们有

Ek−1···E2e1a=AK，

其中，L=Ek−1····E2e1是单位下三角矩阵，而a k（1:k，1:k）是上三角矩阵，因此l a=ak可以写成

，

其中，l1为单位下三角，u1为上三角。（a（1:k，1:k）、l1和u1再次是k×k矩阵；a2和b2是k×n-k矩阵；a3和l3是（n-k）×k矩阵；a4、l4和b4是（n-k）×n-k矩阵。）但是，

l1a（1:k，1:k））=u1，

其中，l1是可逆的（事实上，det（l1）=1），由于假设a（1:k，1:k）是可逆的，所以u1也是可逆的，这意味着（u1）kk=06，因为u1是上三角形。因此，在步骤k中不需要旋转，建立感应步骤。由于det（l1）=1，我们也有det（u1）=det（l1 a（1:k，1:k））=det（l1）det（a（1:k，1:k））=det（a（1:k，1:k）），并且由于u1是上三角形，并且在其对角线上有轴π1，…，πk，我们得到

Det（A（1:K，1:K））=π1π2··································

如要求。

注：如果三角矩阵是可逆的，且其所有对角项都为非零，则可以避免在命题7.2证明的第一部分使用行列式。

推论7.3。（Lu因子分解）设a为可逆n×n矩阵。如果每一个矩阵a（1:k，1:k）对于k=1，…，n都是可逆的，那么高斯消元不需要旋转，得到一个Lu因式分解a=Lu。

证据。我们在命题7.2中证明了在这种情况下，高斯消元不需要旋转。然后，因为每个初等矩阵ei，k；β都是下三角形（因为我们总是把轴πk安排在它操作的行的上方），因为和eks是ei，k；βi，ks的乘积，从

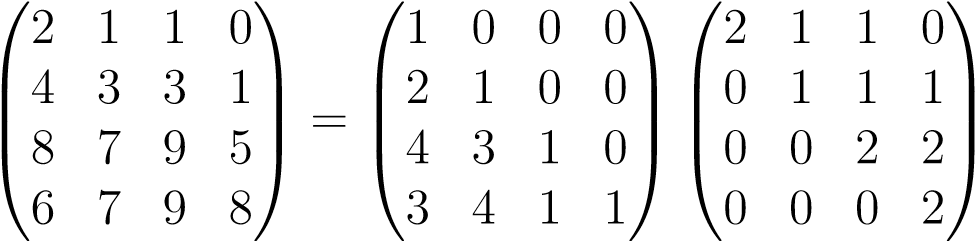
en−1···e2e1a=u，

当u是上三角矩阵时，我们得到

A=lu，

其中是下三角矩阵。此外，由于每个ei，k；β的对角线条目为1，每个ek的对角线条目也为1。

例7.1。读者应该验证



是一个Lu因子分解。

矩阵A的Lu因式分解存在的一个主要原因是，如果我们需要解几个与同一矩阵A对应的线性系统ax=b，我们可以通过解两个三角形系统来实现这一点

Lw=b，Ux=w。

可逆矩阵A的Lu分解a=Lu存在一定的不对称性，实际上，L的对角项都是1，但对于U，这通常是错误的。这种不对称性可以按如下方式消除：如果

D=诊断（U11，U22，…，UNN）

是由u（支点）中的对角线项组成的对角线矩阵，那么我们假设

U0=D−1U，我们可以写

A=ldu0，

其中L为下三角，U0为上三角，L和U0的所有对角线条目均为1，D为支点的对角线矩阵。这样的分解导致以下定义。

定义7.2.我们假设一个可逆的n×n矩阵a如果可以写成a=l d u0，那么它有一个ldu因式分解，其中l是下三角的，u0是上三角的，l和u0的所有对角项都是1，d是对角矩阵。

我们将很快看到，如果a是实对称的，那么u0=l>。

稍后我们将看到，实对称正定矩阵满足命题7.2的条件。因此，对于含有实对称正定矩阵的线性系统，可以采用高斯消元法求解，而无需旋转。实际上，有可能做得更好：这是Cholsky分解。

如果一个平方可逆矩阵A有一个Lu因式分解，那么在进行高斯消元时，可以找到L和U。回想一下，在步骤k，我们选取了一个支点，该支点由矩阵k的第k列的索引j≥k的条目组成，我们交换了行i和k（如果需要）（旋转步骤），然后将索引j=k+1，…，n的条目归零，在k列中，我们得到了以下步骤：

\_\_\_\_\_\_

零

×××

×××。

××××\_

×××。

更准确地说，在排列K行和I行（旋转步骤）之后，如果K行下面的K列中的条目是αk+1K，…，αnk，那么我们将−αjk/πk乘以K行添加到J行；该过程如下所示：

②

所有

……\_\_

A（NKK）

如果我们把j=k+1，…，n写成'jk=αjk/πk，…，n，l的第k列是

.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 网络错误 | | | | |
| 网络错误 |  |  |  |  |
| 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误 |

很容易看出（在定理7.5中证明了这一点），通过翻转“ij:

.

此外，如果高斯消元（无旋转）的结果是u=en−1····e1a，那么

1···························0

…

……………………………………………………………

   

0 1 0 0\_

Ek=0·······································································

   

………………………………………………………………………………

γ

0·······································

所以Ek的kth列是l-1的kth列。

下面是一个例子说明了该方法。

### 例7.2。鉴于

，

−−

我们有以下步骤：第一个支点是第1行的π1=1，我们从第2、3和4行减去第1行。我们得到

.

下一个支点是第2行的π2=−2，我们从第4行减去第2行（并将第2行增加0倍到第3行）。我们得到

.

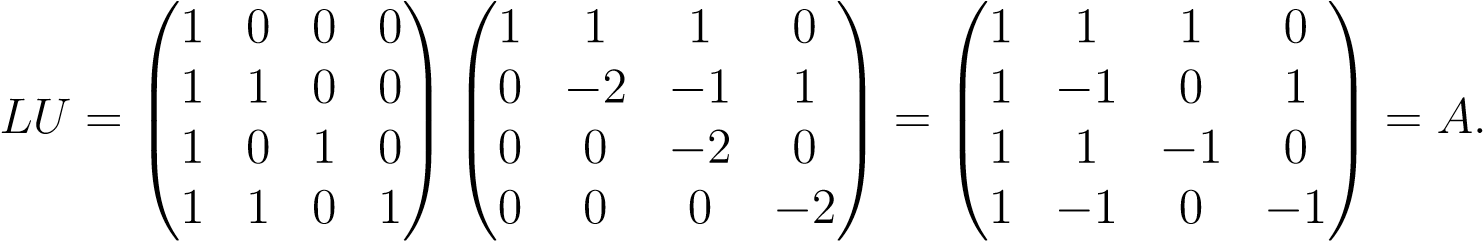
下一个支点是第3行的π3=−2，由于第3列中的第四个条目已经是零，所以我们将第3行的0倍添加到第4行。我们得到

.

程序完成了，我们已经

.

这确实很容易检查



现在我们演示如何扩展上述方法来有效地处理旋转。这是pa=lu因子分解。

## 7.5 pa=lu因子分解

下面的简单命题表明，原则上，a可以由一些置换矩阵p预乘，从而使pa可以在不使用任何支点的情况下转换为上三角形式。排列在第29.3节中有详细讨论，但目前我们只需要这个定义。关于置换概念（如第29.3节所述）与置换矩阵之间的精确联系，请参见问题7.16。

定义7.3.置换矩阵是一个方阵，每行每列只有一个1，其他地方都是零。

如29.3节所示，每个置换矩阵都是换位矩阵（p（i，k）s）的乘积，并且p与逆p>是可逆的。

提案7.4.设a为可逆n×n矩阵。有一些置换矩阵p，所以（pa）（1:k，1:k）对于k=1，…，n是可逆的。

证据。案例n=1是微不足道的，案例n=2也是如此（如果需要，我们交换行）。如果n≥3，我们通过归纳法进行。因为a是可逆的，所以它的列是线性独立的；特别是，它的第一个n-1列也是线性独立的。删除a的最后一列。由于其余n-1列是线性无关的，因此相应n×（n-1）矩阵中也有n-1个线性无关的行。因此，这些n行的排列使得由前n-1行组成的（n-1）×（n-1）矩阵是可逆的。但是有一个对应的置换矩阵p1，因此p1 a的前n-1行和列形成一个可逆矩阵a0。将诱导假设应用于（n-1）×（n-1）矩阵a0，我们发现存在一些置换矩阵p2（保持第n行不变），因此（p2p1a）（1:k，1:k）是可逆的，对于k=1，…，n-1。因为a在第一个位置是可逆的，p1和p2是可逆的，p1p2a也是可逆的，我们就这样做了。

注：我们也可以用Trefethen和Bau[171]提出的高斯消元步骤的巧妙重新排序来证明7.4号命题（第21课）。实际上，我们知道如果a是可逆的，那么就有置换矩阵π和初等矩阵的乘积。

嗯，所以

an=en−1pn−1····e2p2e1p1a，

其中u=an为上三角形。例如，当n=4时，我们有e3p3e2p2e1p1a=u。我们可以定义新的矩阵，这些矩阵仍然是初等矩阵的乘积，因此我们有



事实上，如果我们允许，而且，我们很容易证实

这是初等矩阵的乘积

.

也可以证明是下三角形（见定理7.5）。

一般来说，我们让

ek0=pn−1···pk+1ekpk−+11··pn−11，

我们有

en0−1···e10 pn−1···p1a=u，

其中每个都是一个下三角矩阵（见定理7.5）。

值得注意的是，如果在高斯消元过程中需要旋转步骤，那么对查找Lu因子分解的算法进行一次非常简单的修改就可以得到矩阵l、u和p，从而使p a=Lu。为了描述这个新方法，因为l的对角线项是1s，所以写起来很方便

L=I+∧。

然后，在对矩阵∧进行旋转高斯消元时，我们对∧（真∧k−1）行进行相同的换位，我们在涉及行k和行i的旋转步骤中对A（真的ak）行进行换位。我们也通过从单位矩阵开始对P进行组装。应用于p的行换位与我们应用于a和∧的行换位相同。下面是一个例子来说明这个方法。

### 例7.3。鉴于

，

我们有以下步骤序列：我们初始化∧=0和p0=i4。第一个支点是第1行的π1=1，我们从第2、3和4行减去第1行。我们得到

.

下一个支点是第3行的π2=−2，因此我们对第2行和第3行进行排列；我们还将此排列应用于∧和P：

.

接下来，我们从第4行减去第2行（并将第2行的0倍添加到第3行）。我们得到

.

下一个支点是第3行的π3=−2，由于第3列中的第四个条目已经是零，所以我们将第3行的0倍添加到第4行。我们得到

.

程序完成了，我们已经

.

这确实很容易检查

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 网络错误 |  | 网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误 |  | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 |  | 网络错误  网络错误 | 网络错误  网络错误 |  |
|  |  | 网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 |  | 网络错误  网络错误 | 网络错误  网络错误  网络错误 | 1. 网络错误 2. 网络错误   网络错误  网络错误 |

利用上述例子前面的注释中的思想，我们可以证明下面的定理，用高斯消元的一个简单的适配，证明了计算P、L和U的算法的正确性。

在标准文本中，我们不知道定理7.5的详细证明。尽管Golub和van Loan[80]将这个定理的一个版本表述为他们的定理3.1.4，但他们说“证明是一个混乱的订阅论点。”Meyer[122]还提供了证明的草图（见第3.10节的结尾）。鉴于这种情况，我们提供了一个完整的证据。它确实涉及许多下标和上标，但在我们看来，它包含了一些远远超出符号操作的技术。

定理7.5。对于每个可逆n×n矩阵a，以下保持：

1. 有一些置换矩阵p，一些上三角矩阵u和一些单位下三角矩阵l，所以pa=lu（回忆一下，i=1，…，n时lii=1）。此外，如果p=i，那么l和u是唯一的，它们是由

无旋转的高斯消元。

1. 如果en−1…e1a=u是高斯消去的结果，而没有旋转，那么照常写ak=ek−1…e1a（with），并让1≤k≤n−1和k+1≤i≤n。然后

，

其中，L的第k列是Ek−1的第k列，对于k=1，…，N−1。

1. 写下如果en−a1pk n=−1e·······k−1ep1kp−11a············································

ejj=ej

ejk=pkejk−1pk，对于k=j+1，…，n−1，

和

.

然后，

ejk=pk pk−1···pj+1ejpj+1···pk−1pk u=enn−11··e1n−1pn−1···p1a，

如果我们设置

，

然后

PA=Lu.（†1）

此外，

，

其中ejk是形式的下三角矩阵

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 |

，

和

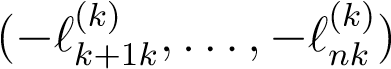
ej=pkejk−1，1≤j≤n−2，j+1≤k≤n−1，k

其中，对于某些i，p k=i或pk=p（k，i），例如k+1≤i≤n；如果pk=6i，这意味着这是通过排列j列中i和k行上的条目获得的。因为矩阵（ejk）−1都是下三角，矩阵l也是下三角。

为了找到l，定义形式的下三角n×n矩阵∧k

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 |

按如下迭代方式组装l的列：let



是Ek第k列的最后一个n−k元素，并通过设置感应地定义∧k

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误 |  |  |  |  |
| 网络错误  网络错误 |  |  |  |  | 网络错误 |
| 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  l0 k0（k-1）1）alk（+1（k…1）k-1）  lnk0（k-1）1 | o  零  零  （k k+1）k  是啊。  （无具体状态） | （b）  是啊。  是啊。  是啊。  （b）  是啊。  是啊。  （b） | -什么？  是啊。  是啊。  （b）  -什么？  （b） | 0-63734；  0-63735；  63735；  0-63735；-63735；  …63735；-63735；-63735；  0-63735；-63735；，  63735；  0-63735；-63735；  …63735；-63735；-63736；  o |

对于某些i>k，pk=i或pk=p（k，i）。这意味着在装配l时，需要对∧k-1的k行和i行进行排列，此时需要对ak的k行和i行进行旋转步骤排列。然后

i+∧k=（e1k）−1···（ekk）−1∧k=e1k+····+ekk，

对于k=1，…，n-1，因此

L=I+∧N−1。

定理7.5的证明是非常技术性的，在第7.6节中给出。

我们再次强调，定理7.5的第（3）部分显示了一个显著的事实，即在对矩阵L进行旋转高斯消元的同时，对算法的唯一更改是对∧k−1行进行相同的换位，我们在涉及行k和行i的旋转步骤。我们也可以通过从单位矩阵开始，并将我们应用于a和∧的相同行转置应用于p来组装p。下面是一个例子来说明这个方法。

例7.4。考虑矩阵

.

我们设置p0=i4，也可以设置∧0=0。第一步是使用轴4排列第1行和第2行。我们还将此排列应用于p0：

.

接下来，我们从第2行减去第1行的1/4倍，从第3行减去第1行的1/2倍，然后将第1行的3/4倍加到第4行，开始装配∧：

.

接下来，我们使用轴5排列第2行和第4行。我们还将此排列应用于∧和p：

.

接下来，我们将第2行的1/5倍增加到第3行，并更新∧02：

.

接下来，我们使用轴6排列第3行和第4行。我们还将此排列应用于∧和p：

.

最后，我们从第4行减去第3行的1/3倍，并更新∧03：

.

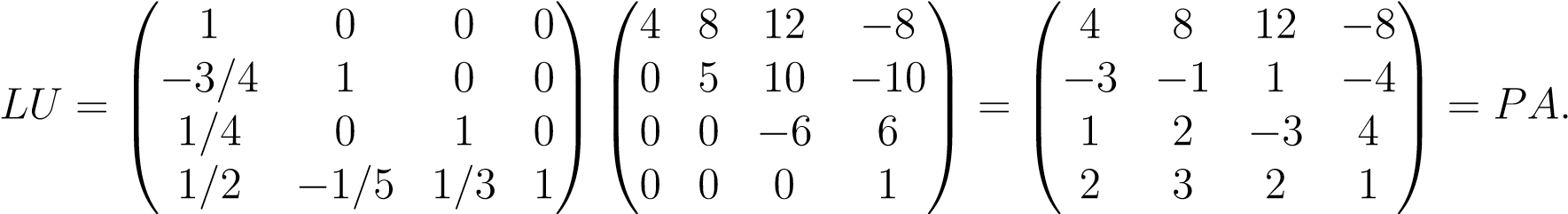
因此，加上∧3的恒等式，我们得到

.

我们检查一下

，

还有那个



注意，如果你愿意覆盖进化矩阵A的下三角部分，你可以在那里存储进化∧，因为这些条目最终将是零！也不需要显式地保存置换矩阵p。您可以将置换步骤记录在一个额外的列中（记录与应用于行的置换π对应的向量（π（1），…，π（n））。我们让读者写了这样一个大胆而有空间效率的版本的吕分解！

**Remark:** In Matlab the function lu returns the matrices *P,L,U* involved in the *PA* = *LU* factorization using the call [*L,U,P*] = lu(*A*).

As a corollary of Theorem 7.5(1), we can show the following result.

**Proposition 7.6.** *If an invertible real symmetric matrix A has an LU-decomposition, then*

*A has a factorization of the form A* = *LDL*>*,*

*where L is a lower-triangular matrix whose diagonal entries are equal to* 1*, and where D consists of the pivots. Furthermore, such a decomposition is unique.*

*7.6. PROOF OF THEOREM 7.5* ~

*Proof.* If *A* has an *LU*-factorization, then it has an *LDU* factorization

*A* = *LDU,*

where *L* is lower-triangular, *U* is upper-triangular, and the diagonal entries of both *L* and *U* are equal to 1. Since *A* is symmetric, we have

*LDU* = *A* = *A*> = *U*>*DL*>*,*

with *U*> lower-triangular and *DL*> upper-triangular. By the uniqueness of *LU*-factorization (Part (1) of Theorem 7.5), we must have *L* = *U*> (and *DU* = *DL*>), thus *U* = *L*>, as claimed.

**Remark:** It can be shown that Gaussian elimination plus back-substitution requires *n*3*/*3+ *O*(*n*2) additions, *n*3*/*3 + *O*(*n*2) multiplications and *n*2*/*2 + *O*(*n*) divisions.

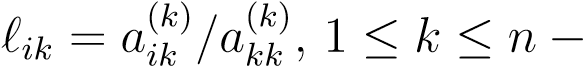
## 7.6 Proof of Theorem 7.5 ~

*Proof.* (1) The only part that has not been proven is the uniqueness part (when *P* = *I*). Assume that *A* is invertible and that *A* = *L*1*U*1 = *L*2*U*2, with *L*1*,L*2 unit lower-triangular and *U*1*,U*2 upper-triangular. Then we have

*L*−2 1*L*1 = *U*2*U*1−1*.*

However, it is obvious that *L*−2 1 is lower-triangular and that *U*1−1 is upper-triangular, and so *L*−2 1*L*1 is lower-triangular and *U*2*U*1−1 is upper-triangular. Since the diagonal entries of *L*1 and *L*2 are 1, the above equality is only possible if *U*2*U*1−1 = *I*, that is, *U*1 = *U*2, and so *L*1 = *L*2.

(2) When *P* = *I*, we have , where *Ek* is the product of *n* − *k* elementary matrices of the form *Ei,k*;−*`i*, where *Ei,k*;−*`i* subtracts *`i* times row *k* from row *i*,

with 1, and *k* + 1 ≤ *i* ≤ *n*. Then it is immediately verified that

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 63723；1  是啊。  63724s；  63724s；  637240；  ek=“63724；637240；0”  63724s；  63744；…  63725秒；  o | -什么？  （b）  -什么？  （b） | o  是啊。  一  -K+1K  是啊。  （b） | o  是啊。  o  一  是啊。  o | -什么？  （b）  -什么？  （b） | 0-63734；  …63735？  63735；  0-63735；  63735；  0'63735；…'63735；'63735；'63736；  一 |