还有那个

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

如果我们把LK定义为

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

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

Lk=Lk−1ek−1，2≤k≤n−1，

因为右边乘以Ek−1会将‘i乘以第i列添加到第k列（矩阵Lk−1的），i>k，而Lk−1的第i列只有非零项1作为其第i个元素。

自从

Lk=e1−1···ek−1，1≤k≤n−1，

我们得出结论，L=ln−1，证明了我们关于L形状的主张。

（3）

第1步。证明（†1）。

首先，我们通过对k的归纳来证明

.

对于k=1，我们有，因为，所以我们的断言无关紧要。

如果k≥2，

AK+1=EKPAK，

根据归纳假设，



7.6。定理7.5的证明~

因为pk是标识或换位，pk2=i，所以通过插入

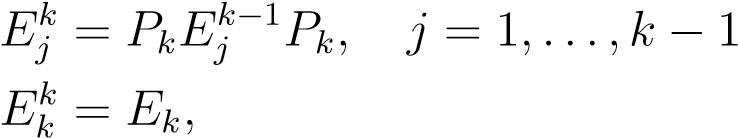
pkpk如下所示，我们可以写

AK+1=EKPAK

=EKPEKK−11···E2K−1E1K−1PK−1···P1A

=ek pkekk−11（pk pk）···（pkpk）e2k−1（pkpk）e1k−1（pkpk）pk−1···p1a=ek（pkkk−11pk）··（pke2k−1pk）（pke1k−1pk）pkpk−1···p1a。

|  |  |
| --- | --- |
| 网络错误 | 网络错误 |



所以我们建立了归纳假设。对于k=n-2，我们得到



如权利要求所述，因数分解pa=lu

P=PN−1···P1

L=（E1n−1）−1····（ENN−11）−1

很清楚。

第2步。证明矩阵（EJK）−1是下三角形。为了实现这一点，我们证明了矩阵ejk是一个非常特殊形式的严格下三角矩阵。

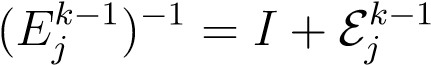
因为对于j=1，…，n-2，我们有ejj=ej，

ejk=pkejk−1pk，k=j+1，…，n−1，

由于pk−1=pk，我们得到（ej j）−1=ej−1，对于j=1，…，n−1，对于j=1，…，n−2，我们得到

.

自从



p k=p（k，i）是一个换位，或者pk=i，所以pk2=i，我们得到

（ejk）−1=pk（ejk−1）−1pk=pk（i+ejk−1）pk=pk2+pk ejk−1 pk=i+pk ejk−1 pk。

因此，我们有

（ejk）−1=i+pk ejk−1 pk，1≤j≤n−2，j+1≤k≤n−1。

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

我们证明了对于j=1，…，n-1，对于k=j，…，n-1，每个都是下三角矩阵，并且

，

对于某些i，p k=i或pk=p（k，i），使得k+1≤i≤n。

对于每个j（1≤j≤n-1），我们通过诱导k=j，…，n-1继续。因为（ejj）−1=ej−1，并且因为ej−1是上述形式，所以基本情况成立。

对于诱导步骤，我们只需要考虑p k=p（k，i）是换位的情况，因为pk=i的情况是微不足道的。我们得弄清楚）是什么。然而，自从

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

因为k+1≤i≤n和j≤k−1，右边的ejk−1乘以p（k，i）将排列列i和k，它们是零列，所以

，

因此，

.

但是自从

（ejk）−1=i+ejk，

我们推断ej=p（k，i）ejk−1。钾

7.6。定理7.5的证明~

我们还知道，将左边的ejk−1乘以p（k，i）将排列行i和k，这表明ejk具有所需的形式，如所述。因为所有EJK都是严格的下三角形

（ejk）−1=i+ejk为下三角形，因此产品



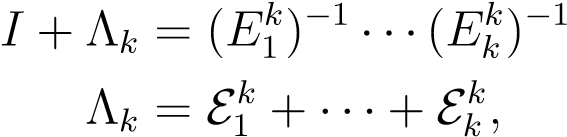
也是下三角形。

第3步。将L表示为L=I+∧n−1，带∧。

从第（3）部分的步骤1，我们知道

.

我们通过对k的归纳证明



对于k=1，…，n-1。

如果k=1，我们有和

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

因为（，我们发现我们得到∧，基本步骤成立。

自（与

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

并且Eikejk=0，如果i<j，如第（2）部分所述，对于涉及Lk的乘积的计算，我们得到

（三）

同样，如果i≥k+1且j≤k−1且自

（ejk）−1=i+pkejk−1，1≤j≤n−2，j+1≤k≤n−1，

我们得到

.（）

通过归纳假设，

，

从（），我们得到

.

使用（），我们推断

.

因为ekk=ek，我们得到

.

但是，根据定义

i+∧k=（i+pk∧k−1）ek−1，

这证明了

，（？）

并完成了验证该公式的诱导步骤。

如果我们用k+1代替k再次应用方程（），我们就得到

，

与（†）一起，我们获得，

，

同时完成了该公式的验证的诱导步骤。对于k=n−1 in（†），我们得到所需方程：l=i+∧n−1。

7.7。处理迂回错误；旋转策略

# 7.7处理迂回错误；关键策略

现在让我们简单地评论一下轴的选择。虽然理论上可以选择任何轴，但舍入误差的可能性意味着选择非常小的轴不是一个好主意。下面的例子说明了这一点。考虑线性系统

10−4x+y=1 x+y=2。

因为10-4是非零的，它可以作为支点，我们得到

10−4X+Y=1

（1−104）Y=2−104。

因此，确切的解决方案是

.

−−

但是，如果四舍五入发生在第四个数字上，则104−1=9999和104−2=9998都将四舍五入到9990，然后解为x=0和y=1，与x≈1和y≈1的精确解相差甚远。问题是我们选了一个非常小的支点。如果我们对方程进行排列，支点是1，在消去之后，我们得到系统。

X+Y=2

（1−10−4）Y=1−2×10−4。

这一次，1−10−4=0.9999和1−2×10−4=0.9998四舍五入为0.999，溶液为x=1，y=1，更接近精确的溶液。

为了解决这个问题，我们可以使用部分旋转策略。这包括在步骤k（1≤k≤n-1）中选择一个条目，以便

.

通过最大化支点的值，我们避免被不希望的小支点分割。注：矩阵a被称为严格的列对角占优iff

，对于j=1，…，n

（响应严格行对角占优iff

，对于i=1，…，n.）

例如，第7.1节中讨论的曲线插值问题的矩阵是严格对角占优的列（和行）。

很长一段时间以来（1900年以前，比如阿达玛），人们都知道，如果矩阵A严格地是列对角占优的（resp）。严格的对角占优行），那么它是可逆的。也可以证明，如果a严格地是对角占优的列，那么

部分旋转的高斯消元实际上不需要旋转（见问题7.12）。

另一种称为完全旋转的策略是选择某个条目，其中k≤i，j≤n，这样

.

然而，在这种方法中，如果所选的轴不在K列中，也需要排列列。这是通过在右边乘以置换矩阵来实现的。然而，在实践中，完全旋转往往过于昂贵，而部分旋转是选择的方法。

LU分解特别有效的一个特殊情况是三对角矩阵的情况，我们现在考虑。

# 7.8三对角矩阵的高斯消去

考虑三对角矩阵

γ

B1

α2

γ

 

A=

γ

γ

γ

γ

γ

定义序列

δ0=1，δ1=b1，

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

δk=bkδk−1−akck−1δk−2，2≤k≤n。

7.8。三对角矩阵的高斯消元

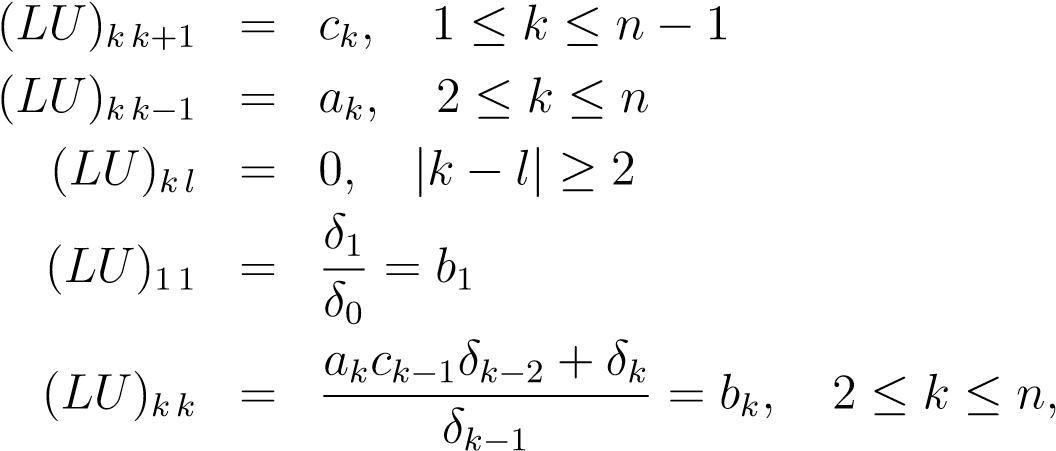
提案7.7.如果a是上面的三对角矩阵，那么k=1，…，n的δk=det（a（1:k，1:k））。

证据。通过将DET（a（1:k，1:k））扩展到其最后一行，该命题随后是对k的归纳。

定理7.8。如果a是上面的三对角矩阵，k=1，…，n的δk=06，则a具有以下Lu因子分解：

.

证据。由于δk=det（a（1:k，1:k））=06，对于k=1，…，n，根据定理7.5（和命题7.2），我们知道a有一个唯一的Lu因式分解。因此，只需检查提议的因子分解是否有效即可。我们很容易查到



因为δk=bkδk−1−akck−1δk−2。

因此，有一个简单的方法来求解线性系统ax=d，其中a是三对角的（而δk=06，k=1，…，n）。为此，可以方便地“挤压”定义的对角矩阵∆使∆k k=δk/δk−1进入因式分解，以便a=（l∆）（∆−1u），并且如果我们允许

，

A=（L∆）（∆−1U）写为

γ

γ

γ

γ

γ

γ

γ

γ

γ

….\_\_

γ

γ

1锌-2\_\_

γ

1锌-1\_\_

γ

一

因此，系统ax=d可以通过构造三个序列来解决：第一，序列

，

−−

对应于递推δk=bkδk−1−akck−1δk−2，通过将该方程的两边除以δk−1得到，下一步

−−

对应于求解系统L∆w=d，最后

xn=wn，xk=wk−zkxk+1，k=n−1，n−2，…，1，

对应于求解系统∆−1ux=w。

备注：可以验证，这需要3（n-1）次加法、3（n-1）次乘法和2n次除法，总共需要8n-6次运算，这比高斯消元所需的O（2n 3/3）要少得多。

我们现在考虑对称正定矩阵（SPD矩阵）的特殊情况。

# 7.9 SPD矩阵和Cholesky分解

回想n×n实对称矩阵a是正定iff

x>ax>0，所有x∈rn，x=06。

等价地，A是对称正定的，如果它的所有特征值都是严格正的。关于对称正定矩阵A的下列事实很容易建立（有些留作练习）：

1. 矩阵A是可逆的。（实际上，如果ax=0，那么x>ax=0，这意味着x=0。）
2. 对于i=1，…，n，我们有aii>0（只要观察x=ei，rn的第i个标准基向量，我们有
3. 对于每一个n×n实可逆矩阵z，矩阵z>az是实对称正定的，如果a是实对称正定的。
4. n×n实对称正定矩阵的集合是凸的。这意味着，如果a和b是两个n×n对称正定矩阵，那么对于任何一个λ∈r，当0≤λ≤1时，矩阵（1−λ）a+λb也是对称正定的。很明显，因为A和B是对称的，（1-λ）A+λB也是对称的。对于任何非零x∈rn，我们有x>ax>0和x>bx>0，所以

x>（（1−λ）a+λb）x=（1−λ）x>ax+λx>bx>0，

因为0≤λ≤1，所以1−λ≥0且λ≥0，且1−λ和λ不能同时为零。

1. n×n实对称正定矩阵的集合是一个圆锥体。这意味着，如果a是对称正定，如果λ>0是任何实数，则λa是对称正定。显然，λa是对称的，对于非零x∈rn，我们有x>ax>0，由于λ>0，我们有x>λax=λx>ax>0。

注：对于复杂的m×n矩阵a，我们将矩阵a定义为m×n矩阵。

A=（艾杰）。然后我们定义a为n×m矩阵a=（a）>=（a>）。如果a=a，n×n复矩阵a是厄米提安矩阵。这是实对称矩阵概念的复杂模拟。厄米矩阵A是正定的如果

z az>0表示所有z∈cn，z=06。

很容易证明厄米正定矩阵的性质（1）-（5）成立；用替换>。

当2×2实对称矩阵A为正定时，它的性质具有指导意义。

写

然后我们有了

.

如果上面的表达式对于所有非零向量都是严格正的，那么对于x=1，y=0我们得到a>0，对于x=0，y=1我们得到b>0。然后我们可以写

.（？）

由于a>0，如果ab−c2≤0，那么我们可以选择y>0，使第二个项为负或零，我们可以设置x=−（c/a）y使第一个项为零，在这种情况下，ax2+2xy+by2≤0，因此我们必须使ab−c2>0。

相反，如果a>0，b>0且ab>c2，则对于任意（x，y）=（06，0），如果y=0，则x=06，且（†）的第一项为正，如果y=06，则（†）的第二项为正。因此，对称矩阵a是正定iff

A>0，B>0，AB>C2.（）

注意ab−c2=det（a），所以第三个条件表示det（a）>0。

观察条件b>0是多余的，因为如果a>0和ab>c2，那么我们必须得到b>0（同样b>0和ab>c2意味着a>0）。

我们可以通过把（a，b，c）看作是x，y，z轴上的坐标，来可视化r3中2×2实对称正定矩阵的空间。由（）中的严格不等式确定的轨迹对应于方程x y=z2圆锥体边上不包含原点且x>0和y>0的区域。当z=δ固定时，方程xy=δ2定义了平面z=δ中的双曲线。方程xy=z2的圆锥体由穿过原点的直线组成，这些直线与平面z=1中的双曲xy=1相接触。我们只考虑这条双曲线的分支，对于它x>0和y>0。见图7.6。

不难证明实对称正定矩阵的逆矩阵也是实对称正定矩阵，但两个实对称正定矩阵的乘积可能不是对称正定矩阵，如下例所示：

.

根据上述准则，左边的两个矩阵是实对称正定的，而右边的矩阵不是均匀对称的，并且

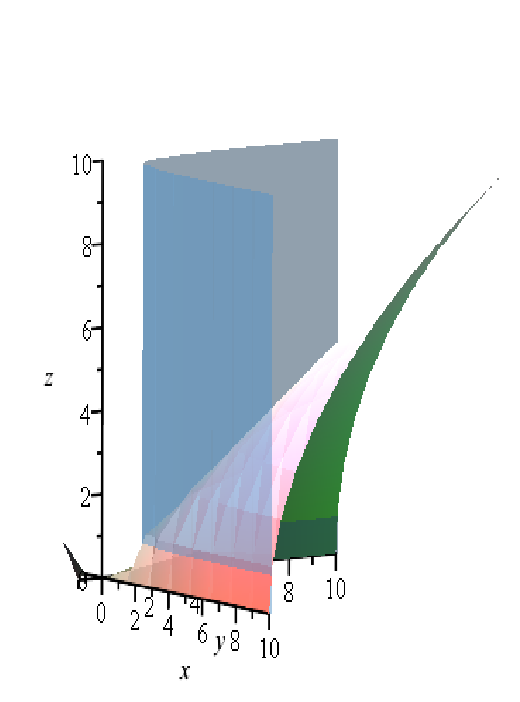
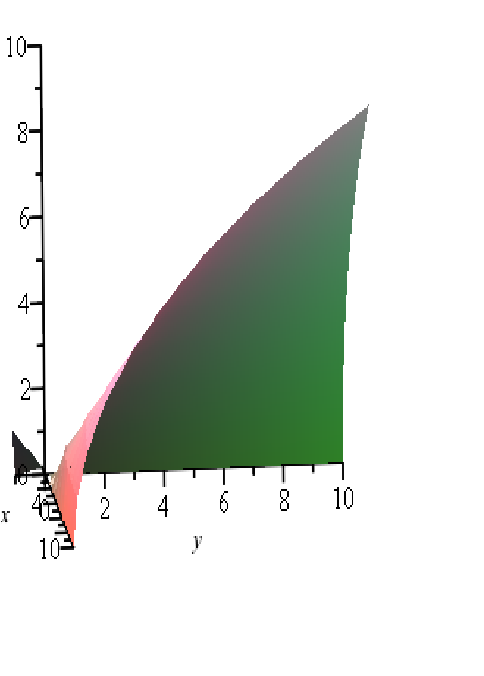
，

即使它的特征值是实的和正的。

接下来，我们证明了实对称正定矩阵具有形式a=b b>的特殊Lu因式分解，其中b是对角元素严格为正的下三角矩阵。这就是乔尔斯基因式分解。

首先，我们注意到一个对称正定矩阵满足命题7.2的条件。

提案7.9.如果a是实对称正定矩阵，那么a（1:k，1:k）是对称正定的，因此k=1，…，n是可逆的。



~~x~~

y

~~=~~

~~1~~

图7.6:r3中表面xy=z2的两个视图。曲面与一个常数z平面的交集产生一条双曲线。与2×2对称正定矩阵相关联的区域位于绿色边的“前面”。

证据。因为a是对称的，所以每个a（1:k，1:k）也是对称的。当W＝Rk，当1个k k为n时，x＝rn为i＝1，…，k，Xi＝0的向量i＝k+1，…，n，因为a是对称正定的，x＝ax＞0，x＝06，x＝06。这尤其适用于从非零向量w∈Rk（如前所定义）获得的所有向量x，并且清楚

x>ax=w>a（1:k，1:k）w，

这意味着a（1:k，1:k）是正定的。因此，根据上面的事实1，a（1:k，1:k）也是可逆的。

命题7.9也适用于复厄米特正定矩阵。命题

7.9可以加强如下：实对称（或复厄米特）矩阵a是正定的iff det（a（1:k，1:k））>0，对于k=1，…，n。

上述事实被称为西尔维斯特准则。我们将在建立Cholsky因子分解之后证明它。

设a为n×n实对称正定矩阵并写出

，

其中c是（n-1）×（n-1）对称矩阵，w是（n√-1）×1矩阵。因为a是对称正定的，a11>0，我们可以计算α=a11。诀窍是我们可以将

，

-

也就是说，作为a=b1a1b1>，其中b1是带有正对角线入口的下三角形。因此，b1是可逆的，根据上面的事实（3），a1也是对称正定的。

备注：矩阵c−ww>/a11被称为矩阵（a11）的schur补码。

定理7.10。（乔尔斯基因式分解）设A为实对称正定矩阵。然后有一些真正的下三角矩阵b，所以a=bb>。此外，可以选择B，使其对角线元素严格为正，在这种情况下，B是唯一的。

证据。我们通过对a的维数n的归纳来进行，对于n=1，我们必须有a11>0，如果我们让α=√a11和b=（α），这个定理就很简单了。如前所述，如果n≥2，我们必须使a11>0，我们可以写

，

-

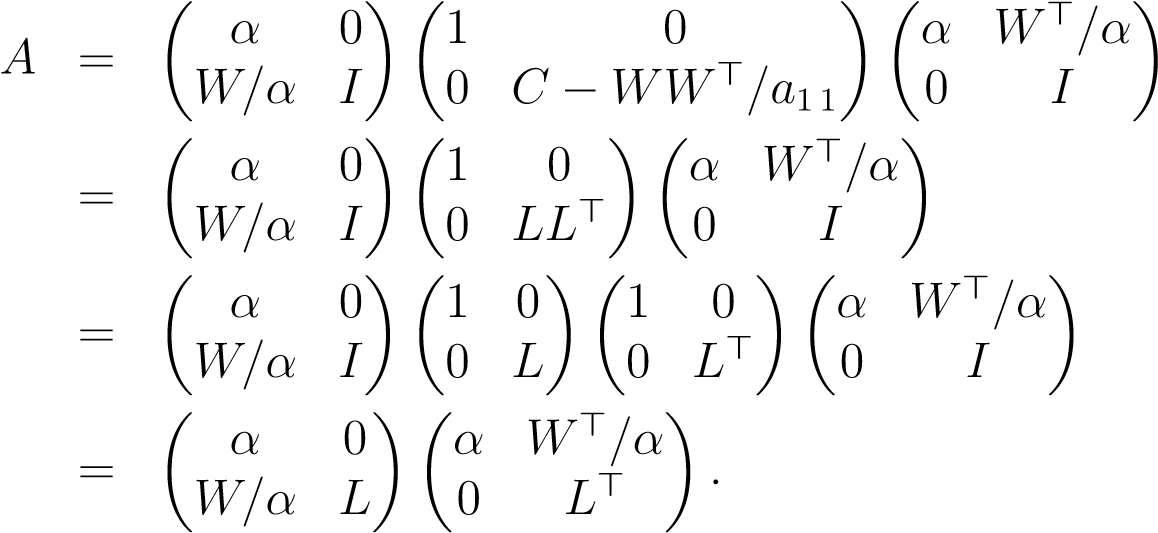
式中α=√a11，矩阵b1可逆，且

-

是对称正定的。然而，这意味着c−ww>/a11也是对称正定的（考虑x>a1x对于x=06和x1=0的每x∈rn）。因此，我们可以将归纳假设应用于c−ww>/a11（这是一个（n−1）×（n−1）矩阵），我们找到一个唯一的下三角矩阵l，其中有正对角项，因此

C−WW>/A11=ll>。

但是我们得到



因此，如果我们

，

我们有一个唯一的下三角矩阵，有正的对角项和a=bb>。

备注：利用一个Lu分解的唯一性，也可以建立Cholesky分解的唯一性。实际上，如果b1和b2是下三角带正对角项，如果我们让∆1（resp.∆2）是由b1的对角线项组成的对角线矩阵。b2）因此（∆k）ii=（bk）ii，对于k=1,2，我们有两个lu分解



对于b1∆−1 1，b2∆−2，1个单位下三角和∆上三角。通过Lu因式分解的唯一性（定理7.5（1）），我们得到

，

第二个方程得出

b1∆1=b2∆2.（）

b1∆1的对角线入口为（b1）2ii，同样地，b2∆2的对角线入口为（b2）2ii，因此上述方程表明：

（b1）2II=（b2）2II，I=1，…，N.

既然b1和b2的对角线入口都是正的，我们必须

（b1）i i=（b2）ii，i=1，…，n；

也就是说，∆1=∆2，由于两者都是可逆的，我们从（）得出结论，b1=b2。

定理7.10也适用于复厄米特正定矩阵。在这种情况下，对于一些具有正对角项的唯一下三角矩阵b，我们有a=bb。

定理7.10的证明立即产生了一种算法，通过求解下三角矩阵b，从a计算b，使a=bb>（其中a和b都是实矩阵）。对于j=1，…，n，

J−1！1/2

bj j=aj j−xb2j k，

K＝1

对于i=j+1，…，n（和j=1，…，n-1）

J−1！

X

bij=aij bikbj k/bj j j.

K＝1

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

上面的公式用于从上到下计算b的jth列，使用之前计算的b的第一个j-1列和矩阵a。在n=3的情况下，a=bb>

B11B31 B21B31+B22B32 B231+B232+B233

我们对a的第一列进行分析，比较条目，并发现

A11=B211 A21=B11B21 A31=B11B31

接下来，我们使用之前计算的B21和B31表达式对A的第二列进行计算，以发现

A22=B221+B222

A32=B21B31+B22B32。

最后，我们使用A的第三列和以前为B31和B32计算的表达式确定B33为

.

例如，如果

我们发现了

我们把它作为一个练习来寻找相似的公式（包括共轭），将一个复厄米正定矩阵A作为a=bb因子。下面的matlab程序实现了cholesky分解。

函数b=cholesky（a）n=大小（a，1）；b=零（n，n）；对于j=1:n-1；如果j=1

b（1,1）=sqrt（a（1,1））；对于i=2:n

b（i，1）=a（i，1）/b（1，1）；结束

其他的

b（j，j）=sqrt（a（j，j）-b（j，1:j-1）\*b（j，1:j-1）'；对于i=j+1:n

b（i，j）=（a（i，j）-b（i，1:j-1）\*b（j，1:j-1）’/b（j，j）；结束

结束

结束

b（n，n）=sqrt（a（n，n）-b（n，1:n-1）\*b（n，1:n-1）'；结束

如果我们在下面的矩阵上运行上述算法

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

Cholesky因式分解可用于求解线性系统a x=b，其中a为对称正定：求解两个系统b w=b和b>x=w。

注：可以看出，该方法需要n3/6+o（n2）加法、n3/6+o（n2）乘法、n2/2+o（n）除法和o（n）平方根提取法。因此，Cholesky方法需要高斯消去所需操作数量的一半（因为高斯消去需要n3/3+o（n2）加法、n3/3+o（n2）乘法和n2/2+o（n）除法）。它还需要一半的空间（只需要B，而不是L和U）。此外，可以证明，Cholesky的方法在数值上是稳定的（见Trefetten和Bau[171]，第23讲）。在matlab中，chol函数返回下三角矩阵b，这样a=bb>使用调用b=chol（a，“lower”）。

注：如果a=b b>，其中b为任意可逆矩阵，则a为对称正定。

证据。显然，b b>是对称的，由于b是可逆的，b>是可逆的，并且

x>ax=x>b b>x=（b>x）>b>x，

很明显，如果x=0.6，x>ax>0

我们现在给出了对称矩阵正定的三个条件。

提案7.11.设a为任意n×n实对称矩阵。以下条件等效：

1. A是正定的。
2. a的所有主要未成年人都是正的；即：Det（a（1:k，1:k））>0，对于k=1，…，n

（西尔维斯特标准）。

1. a有一个lu因子分解，所有的支点都是正的。
2. a有一个ldl>因子分解，d中的所有支点都是正的。

证据。根据命题7.9，如果a是对称正定的，那么每个矩阵a（1:k，1:k）对于k=1，…，n是对称正定的。

k）=q>q，对于某些可逆矩阵q，那么det（a（1:k，1:k））=det（q）2>0。这表明（a）意味着（b）。

如果k=1，…，n的det（a（1:k，1:k））>0，则每个a（1:k，1:k）都是可逆的。通过

命题7.2，矩阵A有一个Lu因式分解，由于支点πk由

γ

如果k=1，则a11=det（a（1:1,1:1））。

如果k=2，…，n，

我们看到πk>0，对于k=1，…，n。因此（b）意味着（c）。

假设A有一个LU因子分解，并且所有的支点都是正的。由于a是对称的，这意味着a具有形式的因式分解

A=低密度脂蛋白>，

l下三角，1在对角线上，其中d是对角线上有正项的对角线矩阵（支点）。这表明（c）意味着（d）。

如果我们形成了对角矩阵，那么在因子分解a=ldl>中，所有的支点都是d正的。

√d=diag（√π1，…，√πn）

γ

如果我们让b=l d，那么我们有

A=bb>，

与B下三角和可逆。根据命题7.11之前的评论，a是肯定的。因此，（d）表示（a）。

准则（c）给出了一个简单的计算测试来检查对称矩阵是否是正定的。对称矩阵是正定的还有一个准则：它的特征值必须是正的。我们必须学习对称矩阵的谱定理来建立这个准则。

命题7.11也适用于复厄米特正定矩阵，在（d）中，因子分解ldl>被ldl替换。

关于高斯消去、Lu因子分解和Cholesky因子分解的稳定性分析和有效实现方法，请参见demmel[49]、Trefethen和Bau[171]、Ciarlet[41]、Golub和van Loan[80]、Meyer[122]、Strang[164、165]和Kincaid和Cheney[100]。

# 7.10缩减行梯队形式（RREF）

第7.2节所述的高斯消元也可应用于矩形矩阵。对于任何矩形m×n矩阵a，这产生了一种确定系统ax=b是否可解的方法和系统可解时所有解的描述。

结果表明，如果我们把所有的支点都重定为1，讨论就简单了，为此我们需要第三类初等矩阵。对于任何λ=06，设ei，λ为n×n对角矩阵

1\_

…

 

1\_

 

ei，λ=λ，

 

1\_

 

…γ

一

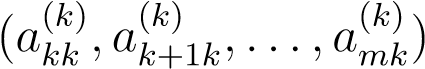
（ei，λ）i i=λ（1≤i≤n）。注意，ei，λ也由下式给出

ei，λ=i+（λ−1）eii，

并且，ei，λ与

ei，λ−1=ei，λ−1.

现在，在K-1消除步骤之后，如果底部



在当前矩阵的第k列中，ak是非零的，因此可以选择一个轴πk，在必要的行排列之后，我们还将行k除以πk以获得轴1，并且不仅使K列中的所有条目i=k+1，…，m归零，而且还使所有条目i=1，…，k−1，so t归零。k列中唯一的非零项是k行中的1。这些行操作是通过将左边的元素矩阵相乘来实现的。

如果=0，则转到K+1列。

当第k列包含一个轴时，将矩阵转换为rref的过程的第k阶段包括以下三个步骤：

0 0 0 0 0××××\_

（k）

0 0 0 aik××0 0××\_

0 0 0××0 0 0××0

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

0 0×1×x x x x x x x x x x\_elim⇒0 0 1×x

0 0=0 0 0××。

××××××××\_

0 0 0 0 0 0 0

0 0×0×0 0×0×0

如果第k列不包含轴，我们只需继续下一列。

结果是，在执行这些消除步骤之后，我们得到了一个特殊形状的矩阵，简称为缩减行梯队矩阵，简称为RREF。

下面是一个例子说明这个过程：从矩阵开始

，

我们执行以下步骤

，

从第2行和第3行中减去第1行；

，

选择轴2和排列第2行和第3行后，将第2行除以2，从第3行减去第2行；

，

第3行除以−1/2后，第1行减去第3行，第2行减去（3/2）×第3行。

很明显，列1、2和4是线性独立的，列3是列1和列2的线性组合，列5是列的线性组合。

1,2，4。

一般来说，导致缩减梯队矩阵的步骤序列不是唯一的。例如，我们可以选择1而不是2作为矩阵A2中的第二个轴。然而，从任何给定矩阵中获得的简化行梯队矩阵是唯一的；也就是说，它不依赖于在简化过程中遵循的步骤序列。这一事实不容易严格证明，但我们稍后会证明。

如果我们想解形式为ax=b的线性方程组，我们将初等行运算同时应用于矩阵a和右侧b。为了方便起见，我们形成了增广矩阵（a，b），即将b作为额外列添加到矩阵中得到的m×n+1矩阵。A.例如，如果

而且，

那么增广矩阵是

.

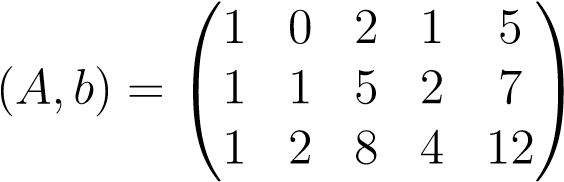
现在对于任何矩阵m，因为

M（A，B）=（MA，MB）

对（a，b）执行基本行操作等同于同时对a和b执行操作。例如，考虑系统

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

它的增广矩阵是矩阵



如上所述，因此应用于该矩阵的还原步骤生成系统

x1+2x3=2 x2+3x3=−1 x4=3。

这个简化后的系统与原系统具有相同的解集，显然x3可以任意选择。因此，我们的系统有无穷多的解

x1=2−2x3，x2=−1−3x3，x4=3，

其中，x3是任意的。

下面的命题表明系统ax=b的解集是由任何行操作序列保留的。

提案7.12。给定任意m×n矩阵a和任意向量b∈rm，对于任意初等行操作序列e1，…，ek，如果p=ek···e1和（a0，b0）=p（a，b），则ax=b的解与a0x=b0的解相同。

证据。因为每个基本行操作ei都是可逆的，所以p也是可逆的，并且（a0，b0）=p（a，b），那么a0=pa和b0=pb。如果x是原始系统ax=b的解，那么两边乘以p，我们得到pax=pb；也就是说，a0x=b0，那么x是新系统的解。相反，假设x是新系统的解，即a0x=b0。

因为a0=pa，b0=pb，p是可逆的，我们得到

ax=p−1a0x=p−1b0=b，

所以x是原始系统的解ax=b。

另一个重要事实是：

提案7.13。给定m×n矩阵a，对于任何行操作序列e1，…，ek，如果p=ek···e1和b=pa，则a行和b行所跨越的子空间是相同的。因此，A和B具有相同的行排名。此外，矩阵A和B也具有相同的（列）秩。

证据。由于b=p a，从先前的观察来看，b的行是a行的线性组合，因此b的行跨度是a行跨度的子空间。由于p是可逆的，a=p−1b，因此通过相同的推理，a的行跨度是a行跨度的子空间。因此，A行和B行所跨越的子空间是相同的，这意味着A和B具有相同的行秩。

命题7.12意味着系统ax=0和bx=0具有相同的解。由于ax是a列的线性组合，bx是b列的线性组合，a中的最大线性独立列数等于b中的最大线性独立列数，即a和b具有相同的秩。

备注：A列和B列所跨越的子空间可以不同！但是，它们的尺寸必须相同。

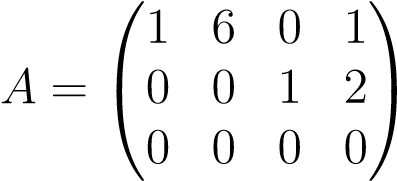
我们将在第7.14节中说明行等级等于列等级。这也将在10.13号提案中得到证明，现在让我们精确地定义什么是缩减的行梯队矩阵。

定义7.4.M×N矩阵A是一个简化的行梯队矩阵，如果满足以下条件：

1. 每行的第一个非零条目是1。此条目称为透视。
2. 第i+1行的第一个非零条目位于第一行的第一个非零条目的右侧。
3. 透视图上方的条目为零。

如果一个矩阵满足上述条件，我们也可以说它是缩减行梯队形式，简称为rref。

请注意，条件（b）意味着透视下的条目也是零。例如，矩阵



是缩减的行梯队矩阵。一般而言，RREF中的矩阵具有以下形状：

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

如果最后一行包含数据透视。

下面的命题表明，使用行操作可以将每个矩阵转换为缩减的行梯队形式。

提案7.14.对于任何m×n矩阵a，都有一个行操作序列e1，…，ek，如果p=ek···e1，那么u=pa是一个缩减的行梯队矩阵。

证据。我们在m上进行归纳。如果m=1，那么这一行中的所有项都是零，因此a=0，或者如果a j是a中的第一个非零项，则让p=（a−j 1）（1×1矩阵）；显然，pa是一个缩减的行梯队矩阵。

现在假设m≥2。如果a=0，我们就这样做了，所以假设a=0.6，因为a=06，有一个最左边的列j是非零的，所以在jth列中选择任何轴π=aij，排列行i和行1，如果需要，将新的第一行乘以π−1，然后通过减去sui清除列j中的其他条目。表1行的倍数。在这个过程的最后，我们有一个矩阵a1，它的形状如下：

，

式中表示任意标量，或更简洁地说

，

7.11。RREF，自由变量，齐次系统

其中d是A（m-1）×N-J矩阵（b是1 J矩阵）。如果j=n，我们就完了。

否则，根据应用于的归纳假设，有一个行操作序列将d转换为缩减的行梯队矩阵r0，并且这些行操作不会影响a1的第一行，这意味着a1被缩减为该形式的矩阵。

.

由于r0是一个约化行梯队矩阵，因此矩阵r满足约化行梯队形式的条件（a）和（b）。最后，通过减去包含一个轴的R0行的适当倍数，可以清除R0中所有轴上的条目。得到的矩阵也满足条件（c），诱导步骤完成。

备注：有一个名为rref的matlab函数，可以将任何矩阵转换为其缩减的行梯队形式。

如果a是任意矩阵，如果r是a的缩减行梯队形式，则命题7.13的第二部分可以稍微尖锐一点，因为缩减行梯队矩阵的结构清楚地表明其秩等于支点数。

提案7.15。矩阵a的秩等于其rref r中的支点数。

# 7.11 RREF、自由变量和齐次线性系统

对于形式为ax=b的系统，我们可以将约简过程应用于增广矩阵（a，b），以获得约简行梯队矩阵（a0，b0），这样系统a0x=b0具有与原始系统ax=b相同的解。约简系统a0x=b0的优点是存在一个si通过简单的测试，检查该系统是否可解，如果可解，则找出其解。

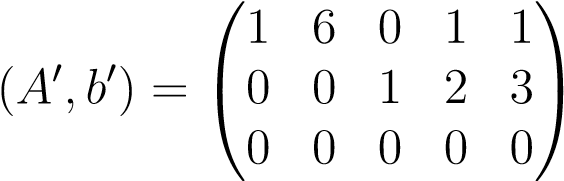
实际上，如果矩阵a0的任何一行为零，并且b0中的相应条目为非零，那么它是一个轴，我们得到了“方程”

0＝1，

这意味着系统a0x=b0没有解决方案。另一方面，如果b0中没有轴，那么对于b0 i=06的每一行i，a0中都有一列j，其中第一行的条目是1（轴）。因此，如果列k不包含轴，我们可以将任意值赋给变量xk，然后求解轴变量。例如，如果我们考虑减少的行梯队矩阵

，

由于第三个方程是0=1，所以没有解a0x=b0。另一方面，简化系统



有解决方案。我们可以任意选取非Pivot列对应的变量x2、x4，然后求解x3（使用第二个方程）和x1（使用第一个方程）。

上述推理证明了以下定理：

定理7.16。如果任何系统ax=b，其中a是m×n矩阵，如果增广矩阵（a，b）是缩减的行梯队矩阵，那么系统ax=b有一个解，如果b中没有轴。在这种情况下，如果列j不包含轴，可以将任意值赋给变量xj。

定义7.5.非透视变量通常称为自由变量。

把命题7.14和定理7.16放在一起，我们得到了一个判定系统ax=b是否有解的标准：将增广系统（a，b）转换成一个行缩减梯队矩阵（a0，b0），并检查b0是否没有支点。

备注：在编写实现行约简的程序时，当到达矩阵A的最后一列时，我们可能会停止。在这种情况下，系统ax=b是否可解的测试是行约简矩阵a0没有索引i>r的零行，这样b0i=06（其中r是支点数，b0是行约简的右侧）。

如果我们有一个齐次系统a x=0，这意味着b=0，当然x=0总是一个解，但是定理7.16意味着如果系统ax=0比方程有更多的变量，那么它有一些非零解（我们称之为非平凡解）。

提案7.17。给定任意一个齐次系统a x=0的m方程在n个变量中，如果m<n，则存在一个非零向量x∈rn，这样ax=0。

证据。将矩阵A转换为缩减的行梯队矩阵a0。我们知道ax=0 iff a0x=0。如果r是a0的支点数，我们必须有r≤m，因此根据定理7.16，我们可以将任意值赋给n-r>0非支点变量，得到非平凡解。

定理7.16也可用于描述平方矩阵可逆的情况。首先，请注意以下简单但重要的事实：

如果一个正方形的n×n矩阵a是一个行约化的梯队矩阵，那么a是一个单位，或者a的底行是零。

提案7.18。设a为尺寸n的平方矩阵。以下条件等效：

7.11。RREF，自由变量，齐次系统

1. 矩阵A可以通过一系列基本的行操作降为单位。
2. 矩阵A是初等矩阵的乘积。
3. 矩阵A是可逆的。
4. 齐次方程组ax=0只有平凡解x=0。

证据。首先，我们证明（a）意味着（b）。如果（a）可以通过一系列行操作e1，…，ep简化为标识，这意味着ep···e1a=i。由于每个ei都是可逆的，我们得到

A=e1−1···ep−1，

其中，每个ei-1也是基本行操作，因此（b）保持不变。现在，如果（b）成立，因为基本行操作是可逆的，那么a是可逆的，（c）成立。如果a是可逆的，我们已经观察到齐次系统a x=0只有平凡解x=0，因为从ax=0，我们得到了−1ax=a−10，也就是x=0。它仍然需要证明（d）意味着（a），为此，我们证明了反面：如果（a）不成立，那么（d）不成立。

利用我们关于降方矩阵的基本观察，如果a不降为恒等式，则a降为一行梯队矩阵a0，其底行为零。假设a0=p a，其中p是基本行操作的产物。因为a0的底行是零，系统a0 x=0最多有n-1个非平凡方程，根据命题7.17，这个系统有一个非平凡解x。但是，ax=p−1ax=0，x=06，这与系统ax=0假设只有平凡解的事实相矛盾。因此，（d）意味着（a）证明是完整的。

命题7.18给出了一种计算可逆矩阵A的逆矩阵的方法：用初等行运算将A简化为恒等式，得到

ep···e1a=i.

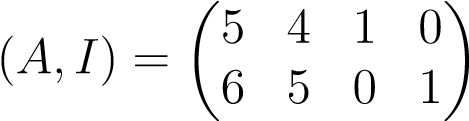
把两边乘以1，我们得到

A−1=ep···e1。

从实际的角度出发，我们可以通过将单位矩阵的n列加到a中得到的增广n×2n矩阵（a，in）简化成行梯队，从而建立产品ep·····e1，这只是执行高斯-乔丹程序的另一种方法。下面是一个例子：让我们找出矩阵的逆矩阵

.

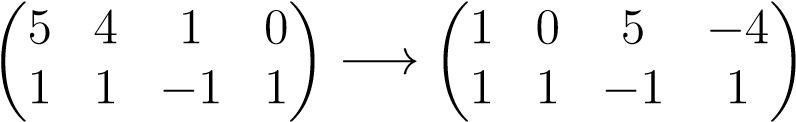
我们形成2×4块矩阵



并应用基本行操作将a约化为标识。例如：



从第2行减去第1行，



从第1行减去4×第2行，

，

从第2行减去第1行。因此

.

命题7.18也可以用来给出一个基本证明，如果一个方阵A有一个左逆B（resp.右倒数b），使b a=i（resp.a b=i），那么a是可逆的，a−1=b。这是一个有趣的练习，试试看！

# 7.12 RREF形式的唯一性

为了完备性，我们证明了矩阵的约化行梯队形式是唯一的。下面给出的整洁的证据是借用并改编自W.Kahan。

提案7.19。设A为任意m×n矩阵。如果u和v是通过应用两个基本行操作序列e1，…，ep和f1，…，fq从a获得的两个缩减行梯队矩阵，那么

u=ep···e1a和v=fq···f1a，

则u=v，ep···e1=fq···f1。换句话说，任何矩阵的缩减行梯队形式都是唯一的。

7.12。RREF的唯一性

证据。让

C=ep····e1f1····fq−1

以便

u=c v和v=c−1u。

我们通过在n上的归纳证明u=v（和c=i）。

设j代表中单位矩阵的第j列，设uj=u`j、vj=v`j、cj=c`j和aj=a`j分别为u、v、c和a的第j列。

首先，我声称uj=0 iff vj=0 iff aj=0。

实际上，如果vj=0，那么（因为u=cv）uj=cvj=0，如果uj=0，那么vj=c−1uj=0。由于u=ep···e1a，我们也得到aj=0 iff uj=0。

因此，我们可以通过从u、v和a中删除由零组成的列来简化我们的任务，因为它们将具有相应的索引。我们仍然用n来表示a的列数。注意，因为u和v是没有零列的缩减行梯队矩阵，所以我们必须有u1=v1=`1。

索赔。如果u和v是没有零列的缩减行梯队矩阵，因此u=cv，对于所有k≥1，如果k≤n，那么‘k出现在u中，iff’k出现在v中，如果‘k确实出现在u中，那么

1. ` k表示U和V中同一列索引jk；
2. U和V的第一个JK列匹配；
3. 索引k+1到m的坐标均等于0的u和v（列索引>jk）中的后续列也匹配。让nk是这样一列的最右边的索引，如果没有，则使用nk=jk。
4. C的第一个NK列与中的第一个NK列匹配。

我们通过对K的归纳来证明这一主张。

对于基本情况k=1，我们已经知道u1=v1=`1。我们也有

c1=c`1=cv1=u1=`1.

如果vj=某些λ∈r的λ'1，则

u j=u`j=cv`j=cvj=λc`1=λc1=λ`1=vj。

使用c−1的一个类似的论点表明，如果uj=λ` 1，那么vj=uj。因此，u和v的所有列与“1”成比例匹配，这就建立了基本情况。注意，如果“2”出现在u中，那么它必须同时出现在u和v中，对于相同的索引，如果不是，则n1=n和u=v。

下一步我们将证明归纳步骤。如果nk=n，那么u=v，我们就完成了。否则，k+1同时出现在u和v中，在这种情况下，通过归纳假设的（2）和（3），它同时出现在u和v中，对于相同的指数，比如jk+1。因此，UJK+1=VJK+1=`K+1。接下来是

c k+1=c`k+1=cvjk+1=ujk+1=`k+1，

因此，c的第一个jk+1列与in的第一个jk+1列匹配。

考虑后面的任何列vj（j>j k+1），其（k+1）之外的元素都将消失。那么vj是vj左边v列的线性组合，所以

uj=cvj=vj。

因为C的前k+1列与中的第一列匹配。同样，任何后面的uj列（j>j k+1），其元素超过（k+1）th都消失，等于vj。因此，u和v（索引>jk+1）中元素超过（k+1）th的所有后续列也都消失匹配，因此c的第一个nk+1列与c的第一个nk+1列匹配，完成了诱导假设。

我们现在可以证明u=v（回想一下，我们可以假设u和v没有零列）。我们之前注意到，u1=v1=`1，所以有一个最大的k≤n，这样‘k出现在u中。那么前面的声明意味着u和v的所有列都匹配，这意味着u=v。

行梯队形式的约简也提供了一种描述形式为ax=b的线性系统解集的方法。

# 7.13使用RREF求解线性系统

首先，我们得到以下简单的结果。

提案7.20。设a为任意m×n矩阵，设b∈rm为任意向量。如果系统

ax=b有一个解，那么这个系统所有解的集合z就是集合

Z=X0+KER（A）=X0+X AX=0，

式中，X0∈Rn是系统a x=b的任意解，即a x0=b（X0称为特殊解），式中，Ker（a）=x∈Rn Ax=0，则是与ax=b相关的齐次系统的解集。

证据。假设系统ax=b是可解的，让x0和x1是任意两个解，这样ax0=b和ax1=b。从第二个方程中减去第一个方程，我们得到

A（x1−x0）=0，

也就是说，x1−x0∈ker（a）。因此，z x0+ker（a），其中x0是ax=b的一个特殊解，反之，如果ax0=b，那么对于任何z∈ker（a），我们有az=0，因此

A（X0+Z）=a x0+a z=b+0=b，

这表明X0+Ker（a）z。因此，z=X0+Ker（a）。

给定一个线性系统ax=b，将增广矩阵（a，b）简化为其行梯队形式（a0，b0）。如前所示，系统ax=b有一个解决方案，如果b0不包含支点。假设是这样。那么，如果（a0，b0）有r个支点，这意味着a0有r个支点，因为b0没有支点，我们知道im的前r列出现在a0中。

我们可以排列a0的列，并相应地在x中重新编号变量，这样im的第一个r列与a0的第一个r列相匹配，然后我们的简化梯队矩阵的形式（r，b0）为

和

其中f是r×（n-r）矩阵，d∈rr。注意r有m-r零行。

那是因为

，

我们看到了

是rx=b0的一个特殊解，因此ax=b。换句话说，我们通过将b0的第一个r分量分配给透视变量，并将非透视变量（自由变量）设置为零得到一个特殊解。

以下是Kumpel Thorpe之前的施工示例。线性系统

x1−x2+x3+x4−2x5=−1

−2x1+2x2−x3+x5=2 x1−x2+2x3+3x4−5x5=−1，

用增广矩阵表示

，

其中a是3×5矩阵。读者应该发现这个系统的行梯队形式是

.

3×5矩阵a0的秩为2。我们排列第二列和第三列（相当于交换变量x2和x3）以形成

.

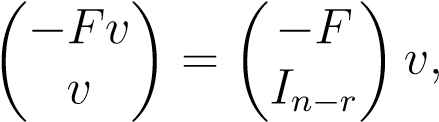
然后给出了这个线性系统的一个特殊解

.

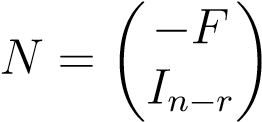
我们还可以找到一个使用f的核（空空间）的基，如果x=（u，v）在a的核中，其中u∈r r和v∈rn−r，那么x也在r的核中，这意味着rx=0；也就是说，

.

因此，u=−fv和ker（a）由形式的所有向量组成。



对于任意的v∈r n−r，它遵循矩阵的n−r列



子矩阵，因此列构成了内核的基础。这是因为线性无关。总之，如果n包含单位矩阵n1，…，n in−n−r ras aa是n的列，那么方程ax=b的一般解由下式给出：

，

其中xr+1，…，xn是自由变量；也就是说，非透视变量。回到上一个例子，我们看到了

，

一般的解决办法是

.

在一般情况下，当轴对应的列与自由变量对应的列混合时，我们发现特殊的解决方案如下。设I1·············<Ir为支点对应列的索引。为k=1，…，r指定透视变量xik，并将所有其他变量设置为0。为了找到核的一个基础，我们形成了n-r向量nk，如下所示。设j1<·······<jn−r为自由变量对应列的指数。对于与自由变量（1≤k≤n−r）相对应的每列jk，形成定义的向量nk，以便条目等于jk列中第一个r条目的负数（翻转这些条目的符号）；设为1，并将所有其他条目设为零。如果索引jk的列（对应于自由变量xjk）是

α1\_

……γ

 

αR\_

，0

 

…γ

零

然后向量nk由

1 0

……γ

I1−1 0 I1−α1 I1+1 0

……γ

IR1 0

IR−αR。

红外+1 0

……\_\_

JK−1 0 JK 1 JK+1 0

………\_\_

N 0

位置jk中的1确保n1，…，nn-r是线性独立的。

作为上述方法的一个例子，考虑找到满足以下性质的n×n矩阵a∈mn（r）的子空间v的基的问题：

1. 每行中的条目之和具有相同的值（例如c1）；
2. 每列中条目的总和具有相同的值（例如c2）。

结果表明，c1=c2，与上述条件对应的2n-2方程是线性无关的。我们把这些事实的证明留作有趣的练习。利用对偶定理（定理10.1）可以证明，满足上述方程的矩阵的空间V的维数是n2−（2n−2）。让我们考虑n=4的情况。有6个方程，空间V有维度10。这些方程是

A11+A12+A13+A14−A21−A22−A23−A24=0 A21+A22+A23+A24−A31−A32−A33−A34=0 A31+A32+A33+A34−A41−A42−A43−A44=0 A11+A21+A41−A12−A22−A42=0 A12+A22+A32+A42−A13−A33−A43=0 A13+A23+A33+A43−A14−A24−A34−A44=0，

相应的矩阵是

1 1 1 1−1−1 0 0 0 0 0 0 0 0 0 0\_

0 0 0 1 1 1 1 1−1−1 0 0 0 0\_

γ

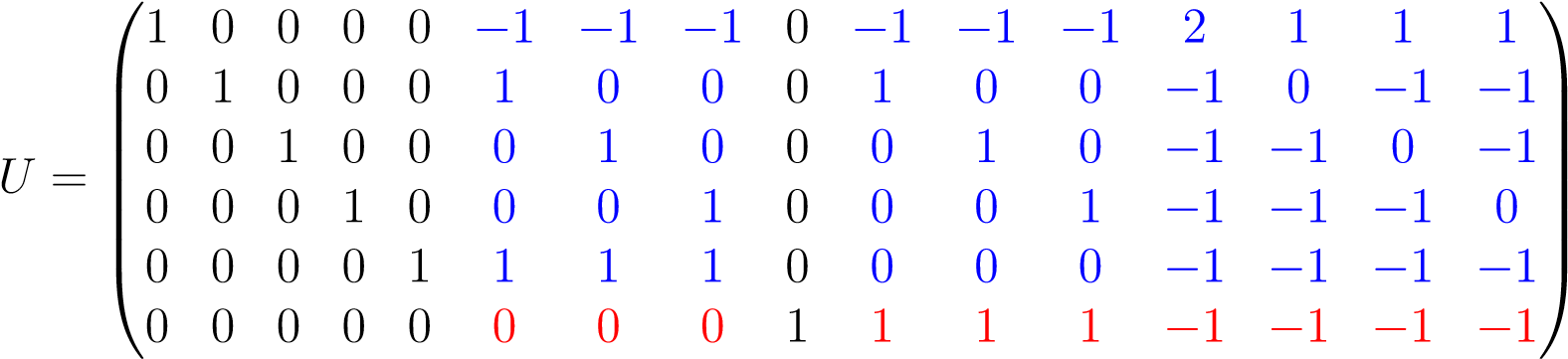
A=01−01 00 00 01−01 00 11−11 01 01−11−01−01。

 

0 1−1 0 0 1−1 0 0 1−1 0 1−1 0\_

0 0 1−1 0 0 1−1 0 0 1−1 0 0 1−1

对行梯队形式进行约简的结果在RREF中生成以下矩阵：



支点变量索引的列表pivlist和自由变量索引的列表free list由下式给出：

pivlist=（1,2,3,4,5,9），freelist=（6,7,8,10,11,12,13,14,15,16）。

应用该算法求出u核的基，得到以下16×10

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

读者应该检查一下，在bk的j列中，最下面的粗体1属于自由列表中索引为jth元素的行，在bk的j列中，索引为pivlist的条目的符号是u列中对应t的6个条目的翻转符号。o自由职业者的第j个指数。现在我们可以从bk中读出构成v基础的4×4矩阵：bk的每一列对应一个已连接行的矩阵。我们得到以下10个矩阵：

1−1 0 1 0−1 0 1 0−1

−1 1 0−1 0 1 0−1 0 0 1

M1=0 0 0，M2=0 0 0，M3=0 0 0，

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

1−1 0 1 0−1 0 1 0−1

0 0 0 0 0 0 0 0 0 0 0

M4=−1 10 0，M5=−1 0 1 0，M6=−1 0 0 1，

1. 0 0 0 0 0 0 0 0 0 0 0 0 0 0

−2 1 1 1−1 0 1 1−1 0 1

m7=11 00 00 00 00，m8=11 00 00 00 00，m9=11 00 00 00 00，

1. 0 0 0 0 1 0 0 0 0 1 0

−1 1 0\_

M10=11 00 00 00 00。

0 0 0 1

回想一下，幻方是一个方阵，它满足关于每一行和每一列中条目之和为同一个数字的两个条件，以及主下降对角线和主上升对角线相加为该公共数字的附加两个约束。此外，这些条目还需要是正整数。对于n=4，附加的两个方程是

A22+A33+A44−A12−A13−A14=0 A41+A32+A23−A11−A12−A13=0，

矩阵为幻方的8个方程是线性无关的。再次，通过行消除，我们得到了“广义幻方”的基础，其条目不限于正整数。我们找到了8个矩阵的基。对于n=3，我们找到3个矩阵的基。

如果一个幻方的项恰好是整数1,2…，n2，则称其为正态。因为这些条目的总和是

，

由于每一行（和每一列）的和都是相同的数字，所以这个公共值（魔法和）是

.

很容易看出n=2没有正规的魔方。对于n=3，魔法和为15，对于n=4，魔法和为34，对于n=5，魔法和为65。

在n=3的情况下，我们有一个附加条件，行和列加起来是15，所以我们得到了一个由两个数字x1，x2参数化的解；也就是说，

.

因此，为了找到一个正规的幻方，我们有额外的不等式约束。

x1+x2>5 x1<10 x2<10 2x1+x2<20

2x1+x2>10 x1>0 x2>0 x1+x2<15，

7.14。初等矩阵和列运算

矩阵中的所有9个条目都必须是不同的。经过一个冗长的案例分析，我们发现了一个显著的事实，那就是有一个独特的正态魔方（直到旋转和反射）：

|  |  |  |
| --- | --- | --- |
| 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 |

结果表明，n=4有880个不同的法向幻方，n=5有275305224个法向幻方（直到旋转和反射）。即使对于n=4，也需要大量的工作才能将它们全部列举出来！找到n>5的魔方数量是一个悬而未决的问题！

# 7.14基本矩阵和柱运算

我们不需要在矩阵A上执行基本的行操作，而是可以执行基本的列操作，这意味着我们用右边的基本矩阵乘以a。作为基本的行和列操作，p（i，k），ei，j；β，ei，λ执行以下操作：

1. 作为行操作，p（i，k）将行i和行k进行排列。
2. 作为列操作，p（i，k）排列列i和列k。
3. p（i，k）的倒数是p（i，k）本身。
4. 作为一个行操作，ei，j；β将β乘以j行加到i行。
5. 作为列操作，ei，j；β将β乘以列i添加到列j（注意索引中的开关）。
6. ei，j的倒数；β为ei，j；−β。
7. 作为行操作，ei，λ将行i乘以λ。
8. 作为列操作，ei，λ将列i乘以λ。
9. ei的倒数，λ为ei，λ−1。

我们可以定义约化柱梯队矩阵的概念，并证明每个矩阵都可以约化为唯一的约化柱梯队形式。对于任意m×n矩阵a，如果我们首先将a转换成它的缩减行梯队形式r，很容易看出我们可以应用基本列操作，将r简化为该形式的矩阵。

，

其中r是支点数（在行缩减过程中获得）。因此，对于每个m×n矩阵a，存在两个初等矩阵序列e1，…，ep和f1，…，fq，这样

.

右边的矩阵称为A的秩正规型，很明显，R是A的秩，作为推论，我们得到了以下重要结果，其证明是直接的。

提案7.21。矩阵A及其转置a>具有相同的秩。

# 7.15运输和扩张~

在这一部分中，我们描述了向量空间e的线性同构，它使某个超平面中的每个向量保持不变。这些映射结果是线性映射，在一些合适的基础上由形式为ei，j；β（transvections）或ei，λ（expansations）的初等矩阵表示。此外，交通量生成SL（e）组，而扩张生成GL（e）组。

设h为e中的任何超平面，并选取一些（非零）向量v∈e，使v/∈h，

e=h kv。

假设f:e→e是一个线性同构，使得f（u）=u代表所有u∈h，而f不是同一性。我们有

f（v）=h+αv，对于一些h∈h和一些α∈k，

当α=06时，因为否则我们会得到f（v）=h=f（h），因为h∈h，与f的注入率（v=6h，因为v/∈h）相矛盾。对于任何x∈e，如果我们写

x=y+tv，对于某些y∈h和一些t∈k，

那么f（x）=f（y）+f（t v）=y+tf（v）=y+th+tαv，

因为αx=αy+tαv，我们得到

F（x）−αx=（1−α）y+th

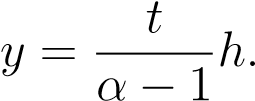
f（x）−x=t（h+（α−1）v）。

如果e是有限维的，通过选择由v和h的基向量组成的e的基，那么f的矩阵是一个下三角矩阵，其对角项都是1，除了第一个等于α的项。因此，Det（f）=α。

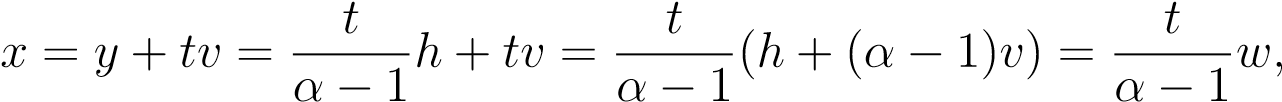
案例1。α=1.6

7.15。转移和扩张~

我们有f（x）=αx iff（1−α）y+th=0 iff



如果我们让w=h+（α-1）v，对于y=（t/（α-1））h，我们有



说明f（x）=αx iff x∈kw。注意w/∈h，因为α=16和v/∈h。

因此，

e=h千瓦，

f是h上的单位和d=kw线上α的放大率。

定义7.6.给定一个向量空间e，对于e中的任意超平面h，任意非零向量u∈e，使u 6∈h，任意标量α=06，1，一个线性映射f，使f（x）=x表示所有x∈h，f（x）=αx表示每个x∈d=ku，称为超平面h，方向d，尺度因子α的扩张。

如果πh和πd是e对h和d的投影，那么我们有

F（x）=πH（x）+απD（x）。

f的倒数由

F−1（x）=πH（x）+α−1πD（x）。

当1时，我们有f2=id，f是关于超平面h方向的对称性。这种情况包括H的正交反射。

案例2。α=1.

在这种情况下，f（x）−x=th，

也就是说，f（x）−x∈kh代表所有x∈e，假设超平面h是作为某种线性形式的核，并设a=\_（v）。我们有a=06，因为v/∈h。对于任何x∈e，我们有\_（x−a−1（x）v）=（x）−a−1（x）（v）=（x）−（x）=0，

这表明x−a−1（x）v∈h代表所有x∈e，由于h中的每一个向量都被f固定，我们得到

X−A−1（X）V=F（X−A−1（X）V）

=f（x）−a−1（x）f（v）

所以

f（x）=x+（x）（f（a−1v）−a−1v）。

由于f（z）−z∈k h对于所有z∈e，我们得出u=f（a−1v）−a−1v=βh对于某些β∈k，所以（u）=0，并且我们有

f（x）=x+（x）u，（u）=0.（）

如上定义的线性图由τ\_，u表示。

相反，对于方程（）给出的任何线性映射f=τ，u，式中，τ是一个非零线性形式，u是一些向量u∈e，因此，如果u=0，则f是恒等式，因此假设u=06。如果是的话，我们有f（x）=x iff（x）=0，也就是说，iff x∈h。我们还声称f的倒数是通过将u改为-u得到的。实际上，我们检查了一个更一般的事实，即

τ，u\_τ，w=τ，u+w.

事实上，利用这个事实，我们有

τ\_，u（τ\_，w（x））=τ\_，w（x）+\_（τ\_，w（x））u

=\_，w（x）+（\_（x）+\_（x）\_（w））u

=τ\_，w（x）+\_（x）u

=x+\_（x）w+\_（x）u=x+\_（x）（u+w）。

对于v=−u，如权利要求所述，我们有τ，u+v=\_，0=id，所以τ，u−1=τ，−u。

因此，我们证明了e的每一个线性同构使某个超平面h中的每一个向量保持不变，并具有f（x）−x∈h（对于所有x∈e）由方程（）定义的映射τ，u给出的性质，式中，ω是定义h的某个非零线性形式，u是h中的某个向量。平均τ\_，u=id iff u=0。

定义7.7.给定e中的任意超平面h，对于任意非零非线性形式，其中，定义h（即h=ker（\_））和任意非零向量u h，线性映射f=τ\_，u（x）=x+（x）u，（u）=0，

对于所有x∈e，称为超平面h和方向u的矢量变换。图f=τ\_，u使h中的每一个矢量保持不变，f（x）−x∈ku表示所有x∈e。

以上参数显示以下结果。

提案7.22。设f:e→e为一个双射线性映射，假设f=6 id，f（x）=x，表示所有x∈h，其中h是e中的某个超平面，如果有一些非零向量u∈e，使得u/∈h和f（u）−u∈h，则f是超平面h的变换；否则，f是超平面的扩张。H.

7.15。转移和扩张~

证据。用上面的符号表示，对于一些v/∈h，我们有f（v）=h+αv，α=06，并写出u=y+tv，y∈h，t=06，因为u/∈h。如果f（u）−u∈h，从

F（U）−U=T（H+（α−1）V）

我们得到（α−1）v∈h，由于v/∈h，我们必须得到α=1，并且我们证明f是一个transvection。否则，α=06，1，我们证明了f是一个膨胀。

如果e是有限维，那么α=Det（f），我们也有以下结果。

提案7.23。设f:e→e为有限维向量空间e的一个双射线性映射，假设f=6 id，且f（x）=x表示所有x∈h，其中h是e中的某个超平面。如果det（f）=1，则f是超平面h的一个矢量化；否则，f是超平面h的一个扩张。

假设f是超平面h和方向u的扩张，并假设det（f）=α=06，1。选取e的基（u，e2，…，en），其中（e2，…，en）是h的基，则f的矩阵形式为α0···0

0 1 0\_

………………，

γ

1. 0 1

它是形式的初等矩阵。相反，很明显，形式ei，α，α=06，1的每一个初等矩阵都是一个扩张。

现在，假设f是超平面h的一个矢量和u∈h的方向，选取一些v/∈h，并选取h的一些基（u，e3，…，en），使（v，u，e3，…，en）是e的基，因为f（v）−v∈ku，f的矩阵是

1. 0···0

α10\_

………………，

0 0···1

它是e2，1；α形式的初等矩阵。相反，很明显，形式ei，j；α（α=0）6的每一个初等矩阵都是一个transvection。

下面的命题是一个有趣的练习，需要很好地掌握基本的行操作ei，j；β；参见问题7.10和7.11。

提案7.24。给定任意可逆n×n矩阵a，有一个矩阵s

，

α=det（a），其中s是形式为ei，j；β的初等矩阵的乘积；也就是说，s是transvections的组成。

令人惊讶的是，每一个交通工具都是两个扩张的组成部分！

提案7.25。如果场k不具有特征2，则超平面h的每个矢量f可以写成f=d2 d1，其中d1、d2是超平面h的扩张，d1的方向可以任意选择。

证据。选取超平面h的扩张d1和比例因子α=06，1。然后，d2=f d−1使h中的每个矢量保持不变，det（d2）=α−1=16。根据命题7.23，线性映射d2是超平面h的扩张，我们得到f=d2 d1，如所述。

注意，在命题7.25中，我们可以选取α=-1；也就是说，超平面H的每一个矢量化都是关于超平面H的两个对称的组成，其中一个对称可以任意选取。

备注：建议7.25只要k=6 0,1。

现在得到以下重要结果。

定理7.26。设e为特征不等于2的场k上的任何有限维向量空间。然后由变换生成sl（e）组，由扩张生成gl（e）组。

证据。考虑任意f∈sl（e），并让a作为其任何基的矩阵。根据命题7.24，有一个矩阵

，

α=Det（a），其中s是形式为ei，j；β的初等矩阵的乘积。由于det（a）=1，我们得到α=1，并且证明了结果。另外，如果f是可逆的，但f/∈sl（e），则上述方程表明en，α是一个扩张，s是一个扩张的乘积，根据命题7.25，每个扩张都是两个扩张的组合。因此，第二个结果也得到了证明。

我们通过证明任意两个矢量在gl（e）中是共轭的来结束这一节。

设τ\_，u（u=0）6为一个矢量，设g∈gl（e）为任意可逆线性映射。我们有

（g\_τ，u\_g−1）（x）=g（g−1（x）+\_（g−1（x））u）=x+（g−1（x））g（u）。

让我们找出由线性形式x→7\_（g−1（x））确定的超平面。这是一组向量x∈e，使得（g−1（x））=0，其中包含iff g−1（x）∈h iff x∈g（h）。因此，Ker（\_g−1）=g（h）=h0，我们得到g（u）∈g（h）=h0，所以gτ，u g−1是超平面h0=g（h）和方向u0=g（u）（带u0∈h0）的变换。

7.16。总结

相反地，让τψ，u0是一些矢量（u0=0）6。选取一些向量v，v0，使\_（v）=ψ（v0）=1，这样

e=h kv=h0 kv0。

存在一个线性映射g∈gl（e），使得g（u）=u0，g（v）=v0，和g（h）=h0。要定义g，选择一个基（v，u，e2，…，en-1），其中（u，e2，…，en-1）是h的基，选择一个基（v0，u0，e02，…，e0n-1），其中（u0，e02，…，en0-1）是h0的基；然后定义g，使g（v）=v0，g（u）=u0，g（ei）=g（e0i），对于i=2，…，n-1。如果n=2，则ei和e0i丢失。那么，我们有了

（g\_τ，u\_g−1）（x）=x+（g−1（x））u0.

现在，\_g−1也决定了超平面h0=g（h），所以我们有\_g−1=λψ用于k中的一些非零λ。由于v0=g（v），我们得到

⑨（v）=⑨g−1（v0）=λψ（v0）、

既然ω（v）=ψ（v0）=1，我们必须有λ=1。接下来是

（g\_τ，u\_g−1）（x）=x+ψ（x）u0=τψ，u0（x）。

总之，我们几乎证明了以下所有部分的结果。

提案7.27。设e为任意有限维向量空间。对于每一个横矢量τ，u（u=06）和每一个线性图g \\τ，u g-1是超平面横矢量τψ，ug（h0（）和方向0 6=0）的横矢量，有一些（u）（即，g \\9702;（τe\\\\（τe\\，u）这样的，g \9702; 9702; 9702\\9702;，u \\\为每一个其他1；其他G−

单词任何两个transvections（=6 id）在gl（e）中是共轭的。此外，如果n≥3，则可以选择上述线性同构g，使g∈sl（e）。

证据。我们只需要证明，如果n≥3，那么对于任何两个交通量τ，u和τψ，u0

（u，u0=0）6，有一些g∈sl（e），例如τψ，u0=gτ，u g−1。和以前一样，我们选择一个基（v，u，e2，…，en-1），其中（u，e2，…，en-1）是h的基，我们选择一个基（

）是h0的基础，我们将g定义为唯一的线性映射，这样

g（v）=v0，g（u）=u0，g（ei）=ei0，对于i=1，…，n-1。但在这种情况下，h和h0=g（h）的维数都至少为2，所以在包括u0在内的h0的任何基中，都有一些独立于u0的基向量，我们可以用这样一种方式重新缩放，即g在两个基上的矩阵具有行列式+1。

# 7.16总结

本章的主要概念和结果如下：

* 不能通过计算行列式来求解（大）线性系统。
* 上三角（下三角）矩阵。
* 反替换法求解（正替换法）。
* 高斯消去。
* 排列行。
* 旋转消除步骤的中心；旋转
* 换位矩阵；初等矩阵。
* 高斯消元定理（定理7.1）。
* 高斯-乔丹因子分解。
* lulu——因式分解（命题因式分解；an7.2存在的充分必要条件）。
* LDU因子分解。
* “pa=lu定理”（定理7.5）。
* 对称矩阵的因式分解。
* 避免小支点：部分支点；完全支点。
* 三对角矩阵的高斯消去。
* 三对角矩阵的Lu因子分解。
* 对称正定矩阵（SPD矩阵）。
* 乔尔斯基因式分解（定理7.10）。
* 对称矩阵为正定的准则；西尔维斯特准则。
* 减少的行梯队形式。
* 把一个矩形矩阵简化成它的行梯队形式。
* 并求出其解，利用行梯队形式的约简来决定系统的特殊解和齐次系统的基ax=b是否可解，

ax=0.

* 魔方。
* 转移和扩张。

# 7.17问题

问题7.1。用高斯消去法求解下列线性系统：

.

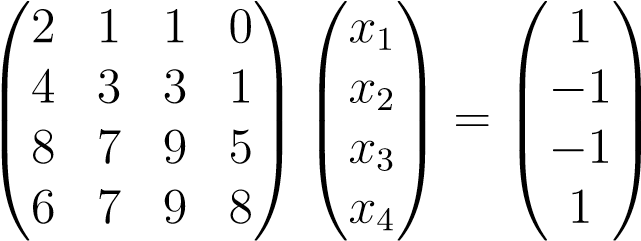
问题7.2。用高斯消去法求解下列线性系统：

.

问题7.3。考虑矩阵

.

当应用高斯消元时，在第二个轴位置，哪个C值为零？在第三个支点位置，C的哪个值为零？在这种情况下，你对矩阵A有什么看法？问题7.4。解决系统问题



使用示例7.1中的lu因子分解。

问题7.5。将rref应用于矩阵

.

问题7.6.将rref应用于矩阵

.

问题7.7。（1）证明2×2矩阵A的子空间的维数，使每一行的条目之和相同（即c1），每一列的条目之和相同（即c2）为2。

1. 证明了2×2矩阵A的子空间的维数，使得每一行的条目之和是相同的（如c1），每一列的条目之和是相同的（如c2），c1=c2也是2。证明每一个这样的矩阵都是

，

并给出这个子空间的基础。

1. 证明了3×3矩阵A的子空间的维数，使得每一行的条目之和相同（如c1），每一列的条目之和相同（如c2），c1=c2为5。首先说明上述约束是由一组方程给出的。

A11\_

A12\_

1 1 1−1−1 0 0 0 A13 0

0 0 1 1 1−1−1 a21 0 1−1 0 1−1 0 a22=0。

γ

   

0 1−1 0 1−1 0 1−1 a23 0

0 1 1−1 0 0−1 0 0 A31 0

 

A32 A33

证明满足上述约束的每一个矩阵都是这样的

，

用a，b，c，d，e∈r.找到这个子空间的基。（使用该方法查找矩阵核心的基础）。

问题7.8。如果a是n×n对称矩阵，b是任意n×n可逆矩阵，证明a是正定的iff b>ab是正定的。

问题7.9。（1）考虑矩阵

.

找到三个形式为e2,1；β1，e3,2；β2，e4,3；β3的矩阵，这样

e4,3；β3e3,2；β2e2,1；β1a4=u4

其中U4是上三角矩阵。计算

m=e4,3；β3e3,2；β2e2,1；β1

检查一下。

.

1. 现在考虑矩阵

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 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 |

找到四个形式为e2,1；β1，e3,2；β2，e4,3；β3，e5,4；β4的矩阵，这样

e5,4；β4e4,3；β3e3,2；β2e2,1；β1a5=u5

其中U5是上三角矩阵。计算

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

m=e5,4；β4e4,3；β3e3,2；β2e2,1；β1

检查一下。

 

0 0 5/4−1\_

0 0 0 0 6/5

1. 编写一个matlab程序，定义函数ematrix（n，i，j，b），它是n×n矩阵，将b乘以j行加到i行。还编写一些matlab代码，生成n×n矩阵，并推广矩阵a4和a5。

用你的程序找出哪五个矩阵ei，j；β把a6减少到上三角矩阵

.

也可以用你的程序计算出哪六个矩阵ei，j；β将a7减少到上三角。-

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 网络错误 |  |  |  |  |  |  |
| 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误 | 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. 求下三角矩阵l6和l7，这样

L6U6=A6

和

L7U7=A7。

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 网络错误 | |  |  |  |  |  |  |
| 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  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. 可以很自然地推测，有n−1矩阵的形式ei，j；β可以减少

也就是说，

e2,1；1/2，e3,2；2/3，e4,3；3/4，···，en，n-1；（n-1）/n.

也可以自然地推测下三角矩阵ln

lnun=安

由给出

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

ln=e2,1；−1/2e3,2；−2/3e4,3；−3/4····en，n−1；−（n−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 | 网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误 |

−

证明上述猜想。

1. 证明a−n1的最后一列是

1/（N+1）

2/（n+1）

……。n/（n+1）问题7.10。（1）设A为任意可逆2×2矩阵

.

证明存在可逆矩阵，这样

，

其中，s是形式ei，j；β的最多四个基本矩阵的乘积。

得出如下结论：sl（2）中的每一个矩阵a（det（a）=+1的可逆2×2矩阵a组）最多是ei，j；β形式的四个基本矩阵的乘积。对于任何a=06，1，给出如上所述的显式因子分解

.

A=-1的分解是什么？

（2）回想一下，旋转矩阵r（群so（2）的成员）是形式的矩阵。

.

证明如果θ=6Kπ（带k∈z），任何旋转矩阵都可以写成一个乘积。

R=ulu，

其中u是上三角形，l是下三角形

.

因此，每个平面旋转（θ=π时围绕原点的翻转除外）都可以写成三个剪切变换的组合！

问题7.11。（1）记住，ei，d是对角矩阵

ei，d=诊断（1，…，1，d，1，…，1）

其对角线项均为+1，但（i，i）第n项等于d。

给定任意n×n矩阵a，对于任意一对（i，j）不同的行指数（1≤i，j≤n），证明存在两个初等矩阵e1（i，j）和e2（i，j），形式为ek，`；β，这样

ej、−1e1（i，j）e2（i，j）e1（i，j）a=p（i，j）a，

通过排列行i和行j从矩阵a中得到的矩阵。等价地，我们得到

e1（i，j）e2（i，j）e1（i，j）a=ej，−1p（i，j）a，

通过排列第i行和第j行并将第j行乘以−1得到的矩阵。

证明对于每一个i=2，…，n，存在四个初等矩阵e3（i，d），e4（i，d），e5（i，d），e6（i，d），形式为ek，`；β，这样

e6（i，d）e5（i，d）e4（i，d）e3（i，d）en，d=ei，d.

当d=-1时会发生什么，也就是说，会发生什么样的简化？

证明了所有置换矩阵都可以写成形式Ek、`；β和形式En、−1的初等运算的乘积。

1. 证明对于每一个可逆n×n矩阵a，都有一个矩阵s

，

d=det（a），其中s是Ek形式的初等矩阵的乘积，`；β。

特别是，sl（n）中的每一个矩阵（det（a）=+1的可逆n×n矩阵A组）都可以写成Ek、`；β形式的初等矩阵的乘积。证明最多需要n（n+1）−2这样的转换。

1. 证明sl（n）中的每一个矩阵最多可以写成（n−1）（max n，3+1）形式的初等矩阵`；β的乘积。

问题7.12。矩阵A称为严格的列对角占优iff

，对于j=1，…，n

证明了如果a是严格的列对角占优，则带部分旋转的高斯消元不需要旋转，a是可逆的。

问题7.13。（1）求下三角矩阵e，使

1 0 0 0 1 0 0 0

E 11 12 01 00=00 11 01 00。

1 3 3 1 0 1 2 1

1. 产品（左边）有什么效果

e4,3；-1e3,2；-1e4,3；-1e2,1；-1e3,2；-1e4,3；-1

在矩阵上

.

1. 求矩阵pa3的逆矩阵。
2. 考虑（n+1）×（n+1）帕斯卡矩阵pan，其第i行由二项式系数给出。

，

与1≤i≤n+1，1≤j≤n+1，并与通常惯例

=0，如果j>i。

矩阵pa3如问题（c）所示，pa4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 |

求n个初等矩阵Eik，jk；βk，这样

.

利用上面的证明，pan的逆矩阵是下三角矩阵，其第i行由有符号的二项式系数给出。

，

1≤i≤n+1，1≤j≤n+1。例如，

10 0 0 0\_

−1 1 0 0 0\_

PA−4 1=1−2 1 0。

 

−1 3−3 1 0\_

1−4 6−4 1

暗示。给定任意n×n矩阵a，将a乘以初等矩阵ei，j；右边的β得到矩阵aei，j；β，其中β乘以第i列加到第j列。

问题7.14。（1）在matlab中实现了将矩形矩阵转换成缩排梯队形式的方法。

1. 用上述方法求可逆N×N矩阵A的逆矩阵，将其应用于通过将单位矩阵的n列添加到a得到的N×2N矩阵[ai]。
2. 考虑矩阵

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 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 |

使用您的程序，找到n=4，…，20的行缩减梯队形式a。

同时运行matlab rref函数并比较结果。

即使对于较小的n值，您的程序也可能与rref不一致。问题在于，某些轴非常小，而规范化步骤（使轴1成为轴1）会导致舍入错误。使用公差参数来解决此问题。

你能推测出A的等级是多少？

1. 证明矩阵A具有以下行约简形式：

1 0−1−2····−（n−2）

0 1 2 3················································

…………………………………

γ

0 0 0 0 0··0

从上面推断A的等级为2。

暗示。一些精心选择的行操作序列。

1. 使用程序显示，如果在a的每个对角线条目中添加大于或等于（2/25）n2的任何数字，将得到可逆矩阵！实际上，运行matlab函数chol应该会告诉您这些矩阵是spd（对称、正定）。

问题7.15。设A为n×n复厄米特正定矩阵。证明了下三角矩阵b的正对角项a=bb由以下公式给出：对于j=1，…，n，

，

对于i=j+1，…，n（和j=1，…，n-1）

J−1！

X

bij=aij bikbj k/bj j j.

K＝1

问题7.16。（排列和排列矩阵）排列可以看作是排列矩阵行的操作。例如，排列

对应于矩阵

.

观察矩阵pπ在每一行和每一列上都有一个1，所有其他项都为零，如果我们将任意4×4矩阵a乘以左边的pπ，则A的行按排列π排列；也就是说，pπa的第（i）行是a的第i行。例如，

.

等价地，pπa的第i行是a的第π−1（i）行。为了使矩阵pπ将a的第i行移动到π（i）第i行，pπ的第i行必须在第一列中有1，在其他地方都有0；这意味着pπ的第i列包含基向量eπ（i），即在π（i）的位置上有一个1，其他地方都有零。

这是一般情况，并导致以下定义。

定义7.8.对于任意置换π：【n】→【n】，表示π的置换矩阵pπ=（pij）是由

；

等价地，pπ的jth列是基向量eπ（j）。置换矩阵p是形式pπ的任何矩阵（其中p是n×n矩阵，π：[n]→[n]是置换，对于某些n≥1）。

注：关于置换矩阵的符号有一个混淆点。置换矩阵p通过对a的行进行置换而在左边乘法作用于矩阵a。如前所述，这意味着pπa的第i行是a的第i行，或者等于pπa的第i行是a的第π−1（i）行。但是观察到中心的行索引第i行的s是π−1（i），而不是π（i）！请参见以下示例：

，

哪里

π−1（1）=4π−1（2）=3π−1（3）=1π−1（4）=2.

证明以下结果

1. 给定任意两个置换π1，π2:[n]→[n]，置换矩阵pπ2\_π1代表π1和π2的组成，等于置换矩阵pπ1和pπ2代表π1和π2的积pπ2pπ1；即，

Pπ2\_π1=Pπ2Pπ1.

1. 表示置换π1的逆矩阵pπ1−1是表示置换π1的矩阵pπ1的逆矩阵pπ−11；也就是说，

.

此外，

Pπ−11=（Pπ1）>。

1. 证明如果p是与换位相关的矩阵，那么det（p）=-1。
2. 证明如果p是一个置换矩阵，那么det（p）=1。
3. 使用置换矩阵给出另一个事实证明，用于表示置换π的置换数的奇偶性仅依赖于π。

244第7章。高斯消元，Lu，Cholsky，梯队形式

第八章

# 向量范数和矩阵范数

## 8.1赋范向量空间

为了定义两个向量或两个矩阵有多接近，并且为了定义向量或矩阵序列的收敛性，我们可以使用范数的概念。回想a，b∈rr+，然后=xz∈=ra−xib≥and0。还记得，如果z=√zz=√a2+zb2=（az+是theib∈cmodulusis a复数，with z）。

定义8.1.设e是一个向量空间，在一个域k上，k是一个域r的实数，或者是复数的域c。e上的范数是一个函数k k:e→r+，将所有x、y、z∈e和λ∈k k u:k的非负实数赋给任意向量u∈e，并满足以下条件

（n1）k x k≥0，kxk=0 iff x=0。（正性）（n2）kλxk=λkxk。（同质性（或标度））。

（n3）kx+yk≤kxk+kyk（三角形不等式）

向量空间e与范数kk一起称为范数向量空间。

通过（n2），设置λ=−1，我们得到

K−XK=K（−1）XK=−1 KXK=KXK；

也就是说，k−xk=kxk。从（n3）开始，我们有

kxk=kx−y+yk≤kx−yk+kyk，

这意味着

KXK−KYK≤KX−YK。

通过交换x和y并利用（n2）这个事实，

KY−XK=K−（X−Y）K=KX−YK，

二百四十五

我们也有

KYK−KXK≤KX−YK。

因此，

| Kxk−Kyk≤Kx−Yk，对于所有x，y∈e.（）

然后，通过设置b，在（n2）中设置λ=0，我们推断k0k=0，而不假设（n1）。y=0 in（），我们得到

| kxk≤kxk，对于所有x∈e。

因此，条件kxk≥0 in（n1）来自（n2）和（n3），并且（n1）可以用较弱的条件代替。

（n1’）对于所有x∈e，如果kxk=0，则x=0，

满足公理（n2）和（n3）的k k:e→r函数称为半范数。由以上讨论可知，半范数对所有x∈e也具有kxk≥0的性质，k0k=0。

然而，可能存在非零向量x∈e，使得kxk=0。

我们来举几个赋范向量空间的例子。

例8.1。

1. 设e=r，kxk=x，x的绝对值。
2. 设e=c，kzk=z，z的模量。
3. 设e=rn（或e=cn）。有三个标准规范。对于每个（x1，…，xn）∈e，我们有规范kxk1，定义如下：

KXK1=x1+········xn，

我们有欧几里得标准kxk2，定义如下：

，

sup-norm kxk∞定义如下：

Kxk＝max＝{Xi}{ 1 } i＝n}。

更一般地说，我们定义“p-norm”（对于p≥1）的方法是

kxkp=（x1 p+······xn p）1/p.

见图8.1至8.4。

K

1

K

0

.

5

0

0

.

5

1

K

1

K

0

.

5

0

.

5

1

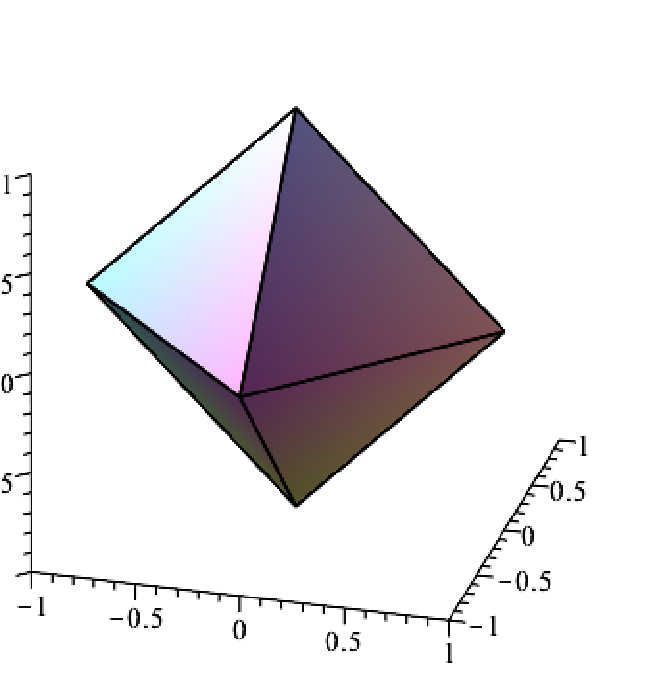
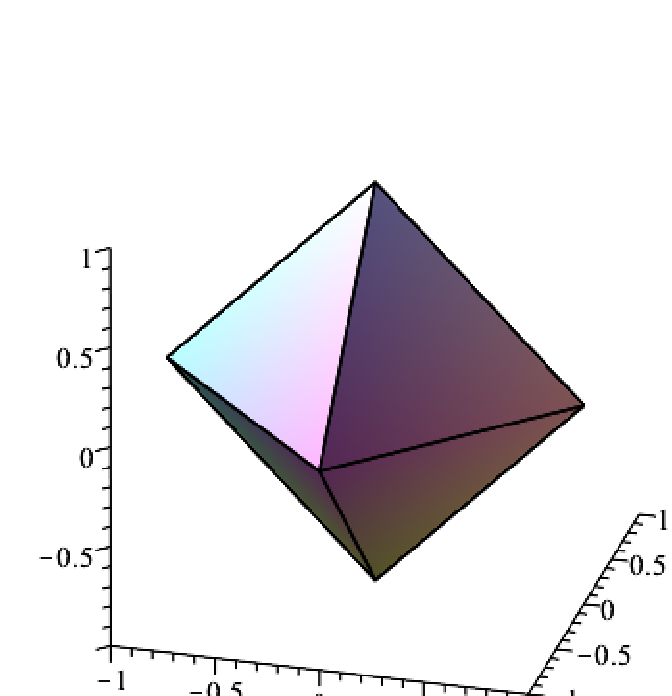


图8.1：上图为x∈r2 kxk1≤1，下图为x∈r3 kxk1≤1。

除了“p-规范”，还有其他规范。下面是一些例子。

1. 对于e=r2，k（u1，u2）k=u1+2 u2。

见图8.5。

1. 对于e=r2，

.

见图8.6。

1. 对于e=c2，k（u1，u2）k=u1+iu2+u1−iu2。

读者应该检查它们是否满足一个规范的所有公理。

需要做一些工作来证明p-范数的三角形不等式。

K

1

K

0

.

5

0

0

.

5

1

K

1

K

0

.

5

0

.

5

1

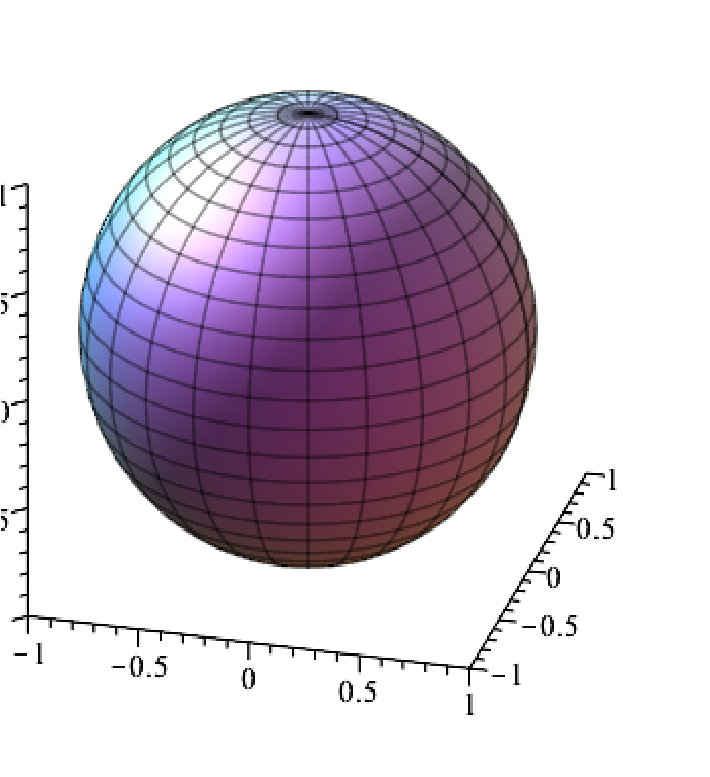
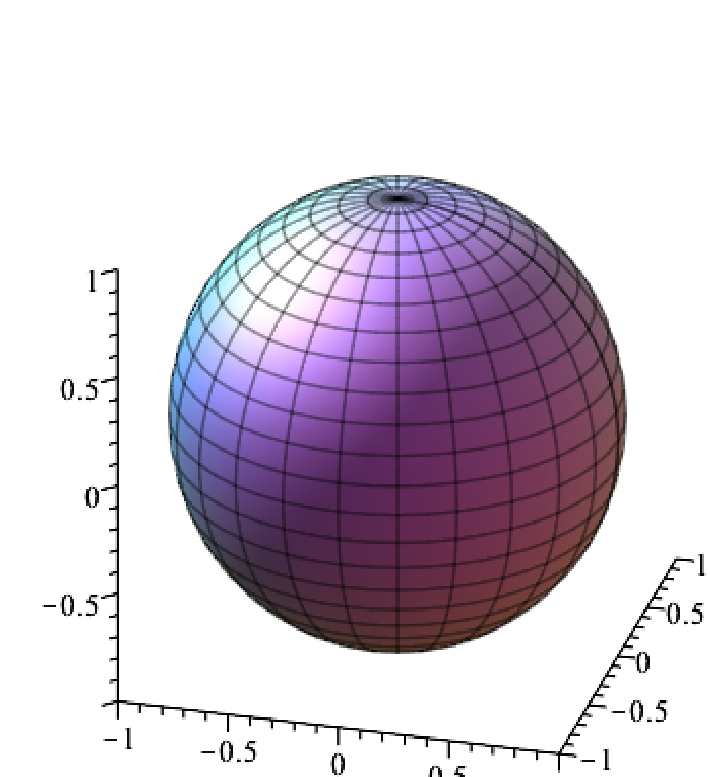


图8.2：上图为xk2≤1。\_x∈r2 kxk2≤1，下图为x∈r3 k

提案8.1.如果e=cn或e=rn，对于每个实数p≥1，'p-范数实际上是一个范数。

证据。案例p=1和p=∞很容易，留给读者。如果p>1，那么让q>1这样

.

我们将利用以下事实：对于所有α，β∈R，如果α，β≥0，那么

.（）

为了证明上述不等式，我们利用指数函数t 7→et满足以下凸不等式的事实：

eθx+（1−θ）y≤θex+（1−θ）ey，

一

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

K1

K1

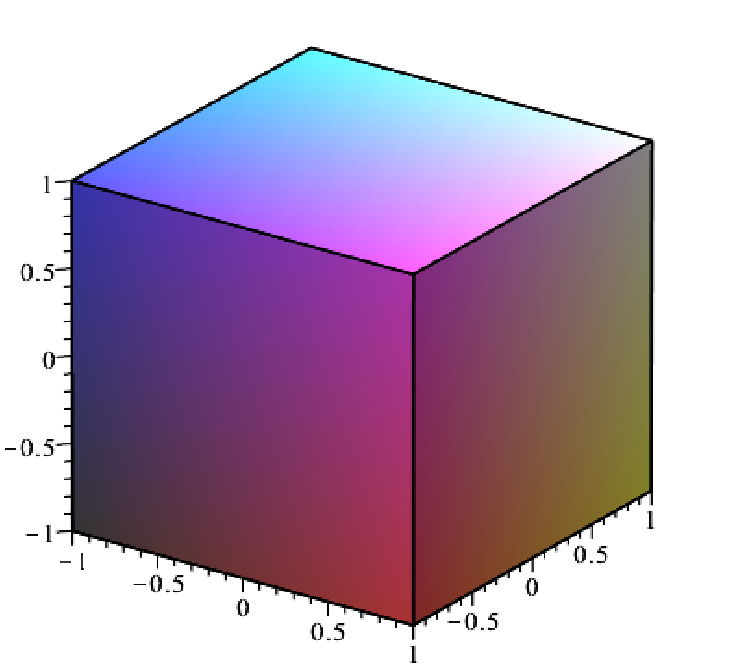
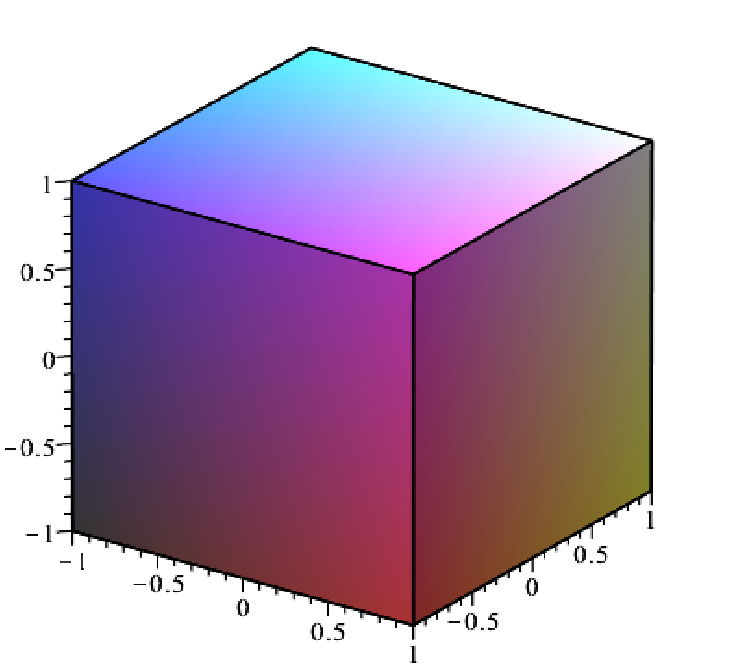


图8.3：上图为x∈r2 kxk∞≤1，下图为x∈r3 kxk∞≤1。

对于所有x，y∈r和0≤θ≤1的所有θ。

由于αβ=0的情况很小，我们假设α>0和β>0。如果我们将θ替换为

1/p，x乘p logα，y乘q对数β，得到

，

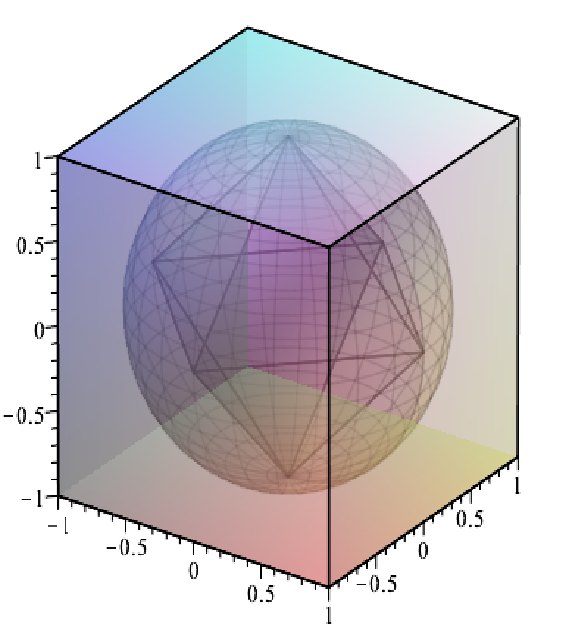
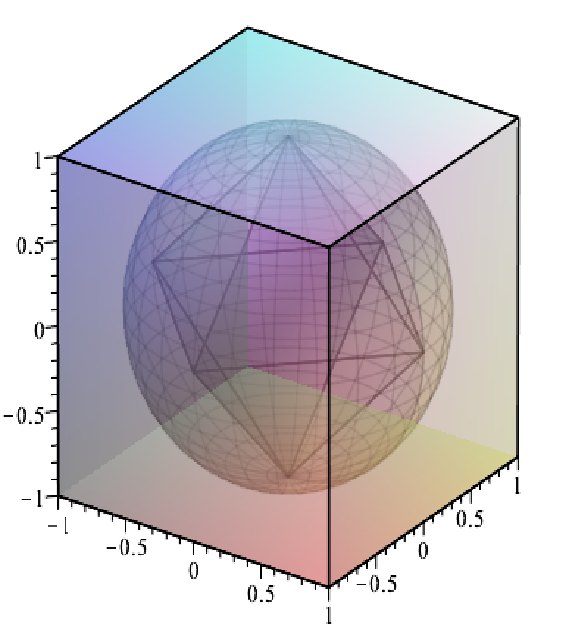
它简化为

，

如要求。

我们现在要证明，对于任意两个向量u，v∈e，（其中e是维度n），我们有

.（）



K

1

K

0.5

0

0.5

1

K

1

K

0.5

0.5

1

图8.4:1-范数、欧几里得范数和sup范数的闭单元球之间的关系。

如果α=ui/uku=0KP和或vβ==0 v的不等式（），由于上述不等式是微不足道的，我们假设i/kvkq产生u 6=0和v=06。然后

，

对于i=1，…，n，通过总结这些不等式，我们得到

，

如要求。为了完成证明，我们只需证明（n3）属性有效，因为（n1）和（n2）是明确的。对于i=1，…，n，我们可以写

（ui+vi）p=ui（ui+vi）p−1+vi（ui+vi）p−1，

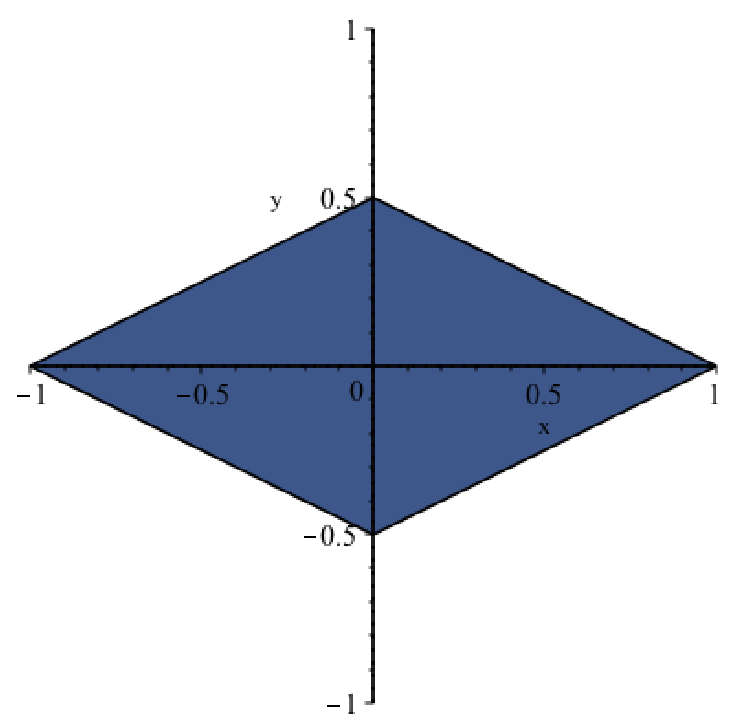


图8.5：单元闭合单元Ballu1+2 u2（u1，u2）∈r2 k（u1，u2）k≤1，其中k（u1，u2）k=

γ

所以通过总结这些方程，我们得到

，

利用不等式（），用v∈e，其中vi=（ui+vi）p−1，我们得到

.

但是，1/p+1/q=1意味着pq=p+q，也就是说，（p−1）q=p，所以我们有

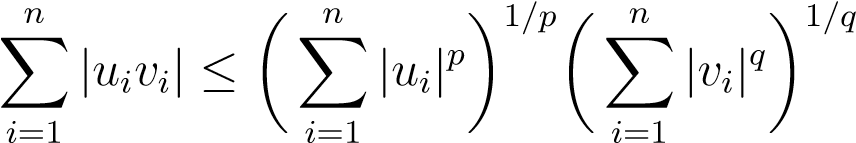
，

会产生

.

由于ui+vi≤ui+vi，以上所述表示三角形不等式ku+vkp≤kukp+kvkp，如权利要求所述。

对于p>1和1/p+1/q=1，不等式



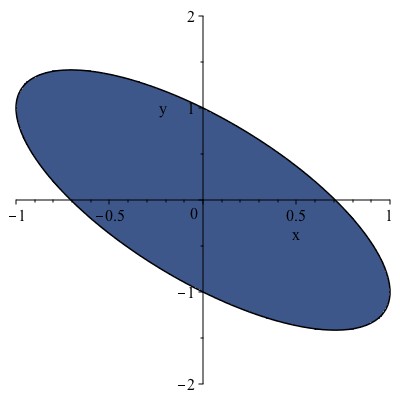


图8.6：单元闭合单元球（u1，u2）∈r2 k（u1，u2）k≤1，其中k（u1，u2）k=

.

被称为H–older不等式。对于p=2，这是柯西-施瓦兹不等式。实际上，如果我们定义Hermitian内积H−，−I on CN by

，

其中u=（u1，…，un）和v=（v1，…，vn），然后

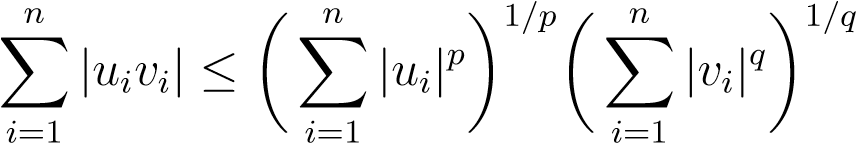
，

所以H–older的不等式意味着以下不等式。

推论8.2。（H——年长者的不等式）对于任何实数p，q，这样p，q≥1和

，

（如果p=1，q=+∞如果q=1，p=+∞）我们有不等式



和

| hu，vi≤kukp kvkq，u，v∈cn.

对于p=2，这是标准的柯西-施瓦兹不等式。p-范数的三角形不等式，

，

被称为闵可夫斯基不平等。

当我们把厄米内积限制为实向量u，v∈rn时，我们得到

欧几里得内积

.

如果我们用列向量表示（通常）u=（u1，…，un）和v=（v1，…，vn）（在rn中），那么它们的欧几里得内积由下式给出：

hu，vi=u>v=v>u，

当u，v∈cn时，它们的厄米田内积由

hu，vi=v u=u v。特别是当u=v时，在复杂情况下，我们得到



在实际情况下，这变成



尽管这些符号很方便，但我们仍然建议您不要滥用它们；符号hu，vi更为固有，当我们的向量空间是无限维时仍然“有效”。

注：如果0<p<1，则x 7→kxkp不是范数，因为三角形不等式失败。例如，考虑x=（2,0）和y=（0,2）。那么x+y=（2,2），我们得到kxkp=（2p+0p）1/p=2，kkp=（0p+2p）1/p=2，kx+ykp=（2p+2p）1/p=2（p+1）/p。

因此，kx+ykp=2（p+1）/p，kxkp+kkkp=4=22。

由于0<p<1，我们得到2p<p+1，即，（p+1）/p>2，所以2（p+1）/p>22=4，三角形不等式kx+ykp≤kxkp+kkkp失效。

观察k（1/2）x k p=（1/2）kxkp=k（1/2）ykp=（1/2）kkkp=1，k（1/2）（x+y）kp=21/p，由于p<1，我们得到21/p>2，所以

k（1/2）（x+y）k p=21/p>2=（1/2）kxkp+（1/2）kkkp，

地图x 7→kxkp不是凸的。

对于p=0，对于任何x∈rn，我们有

kxk0＝{{i {{ 1，…，n}〉Xi 6＝0 }，

x的非零分量的数目。图x 7→kxk0这次不是标准值，因为

AXIOM（N2）失效。例如，

k（1,0）k0=k（10,0）k0=1 6=10=10 k（1,0）k0。

地图x 7→kxk0也不是凸的。例如，

k（1/2）（2,2）k0=k（1,1）k0=2，

和k（2,0）k0=k（0,2）k0=1，

但k（1/2）（2,2）k0=2>1=（1/2）k（2,0）k0+（1/2）k（0,2）k0。

然而，“零范数”x 7→kxk0在机器学习中被用作一个规则化术语，它鼓励稀疏性，即增加向量x的零分量的数量。

下面的建议很容易说明。

提案8.3.以下不等式适用于所有x∈rn（或x∈cn）：

，

命题8.3实际上是一个非常重要结果的特例：在有限维向量空间中，任何两个范数都是等价的。

定义8.2.对于任意（实或复）向量空间e，两个规范k ka和k kb是等价的，如果存在一些正实c1，c2>0，那么

kuka≤c1 kukb，kukb≤c2 kuka，表示所有u∈e。

给定n维向量空间上的任意范数k，对于e的任意基（e1，…，en），观察到对于任意向量x=x1e1+·········+xnen，我们得到kxk=kx1e1+·······························································定义为

kxk1=kx1e1+·····+xnenk=x1········+xn。

上面的意思是

| kuk−kvk≤ku−vk≤c ku−vk1，

这意味着下面的推论。

推论8.4.空间E，任意normu 7→kuk的映射相对于normu 7→kuk在有限维（复杂或真实）vectork k1上是连续的。

设为关于规范k k1的单位球面，即

.

现在是有限维向量空间的一个封闭有界子集，因此由Heine–Borel（或等效地，由Bolzano–Weiertrass）压缩。另一方面，非空紧集上的连续实值函数具有最小值和最大值，这是一个众所周知的分析结果。利用这些事实，我们可以证明以下重要定理：

定理8.5。如果e是有限维的实向量空间或复向量空间，那么e上的任意两个范数都是等价的。

证据。足以证明任何范数k k等于1-范数。我们已经证明了函数x 7→k x k相对于范数k k1是连续的，并且我们观察到单位球面是紧凑的。现在我们回顾一下，因为函数f:x 7→kxk是连续的，并且因为它是紧凑的，所以函数f有一个最小m和一个最大m，因为exk1x=1k永远不会为零，所以我们必须有m>0。因此，我们证明了如果k

0<m≤kxk≤m，

所以对于x 6=0的x∈e，我们得到

m≤kx/kxk1k≤m，

这意味着

m kxk1≤kxk≤m kxk1。

由于上述不等式具有无足轻重的生命等效性，如所声称的.x=0，我们刚刚证明了k k和k k1是

注：P为N×N对称正定矩阵。立即确认地图x 7→kxkp由

kxkp=（x>px）1/2

是RN上的范数，称为二次范数。通过一些凸分析（L–owner–john椭球体），可以证明RN上的任何范数k k都可以近似为二次范数，即存在二次范数k kp，从而

kxkp≤kxk≤√nkxkp，所有x∈rn；

见Boyd和Vandenberghe[29]第8.4.1节。

接下来我们将讨论矩阵上的规范。

## 8.2矩阵规范

为了便于解释，我们将考虑平方n×nnmatries的向量空间mn（r）和mn（c）。大多数结果也适用于空格。因为n×n矩阵可以相乘，所以矩阵范数ism，n（r）和m m，n（c）的思想是，矩形m×的矩阵范数ism，n（r）和mm，n（c）在矩阵相乘方面应该表现得“好”。

定义8.3.方阵n×n矩阵空间上的矩阵范数k k k（k），k=r或k=c，是向量空间mn（k）上的范数，其附加性质称为次多积性，即kabk≤kakkbk，

对于所有a，b∈mn（k）。满足上述性质的矩阵上的范数通常称为子乘法矩阵范数。

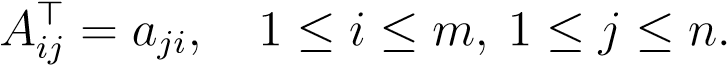
由于i2=i，从kik=ki2k≤kik2，我们得到了每个矩阵范数的kik≥1。

在给出矩阵规范的例子之前，我们需要回顾一些关于

矩阵。给定任意矩阵a=（aij）∈mm，n（c），a的共轭a是这样的矩阵：

aij=aij，1≤i≤m，1≤j≤n。

a的转置是n×m矩阵a>这样



a的伴随是n×m矩阵a，因此

A=（A>）=（A）>。

当a是实矩阵时，a=a>。矩阵a∈mn（c）是厄米提安如果

A=A.

8.2。矩阵范数

如果a是实矩阵（a∈mn（r）），我们认为a是对称的，如果

A>=A。

矩阵a∈mn（c）是正态的，如果

a a=a a，

如果A是一个实矩阵，它是正常的，如果

a a>=a>a。

矩阵u∈mn（c）是一元如果

u u=u u=i.

实矩阵q∈mn（r）是正交的，如果

q q>=q>q=i。

给定任意矩阵a=（aij）∈mn（c），a的迹Tr（a）是其对角元素Tr（a）=a11+······+ann的和。

很容易显示出轨迹是线性图，因此

Tr（λa）=λTr（a）

和Tr（a+b）=Tr（a）+Tr（b）。

此外，如果a是m×n矩阵，b是n×m矩阵，则不难证明

tr（ab）=tr（ba）。

我们还回顾了特征值和特征向量。我们满足于关于矩阵的定义。稍后将进行更全面的治疗（见第14章）。

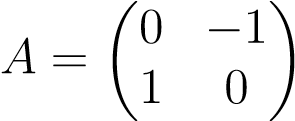
定义8.4.给定任意一个方阵a∈mn（c），如果有一些非零向量u∈cn，则复数λ∈c是a的特征值，这样

au=λu.

如果λ是a的特征值，则非零向量u∈cn使au=λu称为与λ关联的a的特征向量；与零向量一起，这些特征向量形成eλ（a）表示的cn的子空间，并称为与λ关联的特征空间。

注：注意定义8.4要求特征向量非零。这一要求的一个有点不幸的结果是，由于零向量丢失，特征向量集不是子空间！在积极方面，只要涉及特征向量，就不需要说它们是非零的。特征向量是非零的事实在所有涉及它们的论点中都被隐式地使用，因此，规定特征向量应该是非零似乎更安全（但也许不是很优雅）。

如果a是一个平方实矩阵a∈mn（r），那么我们将定义8.4限定为实特征值λ∈r和实特征向量。然而，需要注意的是，虽然每个复矩阵总是至少有一些复特征值，但一个实矩阵可能没有任何实特征值。例如，矩阵



具有复杂特征值i和−i，但没有实际特征值。因此，即使对于实矩阵，我们通常也考虑复特征值。

观察到λ∈c是a的特征值

* 非零向量u∈cn的iff au=λu
* iff（λi−a）u=0
* 如果矩阵λi−a定义了一个具有非零核的线性映射，即，
* iffλi−a不可逆。

然而，根据命题6.10，λi−a不是可逆的iff

Det（λi−a）=0。

现在，det（λi−a）实际上是形式不确定的λ中n次多项式。

λn−tr（a）λn−1+····+（−1）n det（a）。

因此，我们看到a的特征值是上述多项式的零（也称为根）。因为n次的每一个复多项式都有n个根，用它们的多重性来计算，所以我们有以下定义：

定义8.5.给定任意平方n×n矩阵a∈mn（c），多项式

det（λi−a）=λn−tr（a）λn−1+····+（−1）n det（a）

被称为a的特征多项式。特征多项式的n（不一定是不同的）根λ1，…，λn都是a的特征值，构成a的谱。

ρ（a）=maxλi|

1≤I≤N

是A特征值的最大模，称为A的谱半径。

8.2。矩阵范数

由于特征值λ1，…，a的λn是多项式的零。

det（λi−a）=λn−tr（a）λn−1+·····+（−1）n det（a）、

我们推断（详情见第14.1节）

Tr（a）=λ1+·····+λn Det（a）=λ1···························

提案8.6.对于mn（c）上的任意矩阵范数k k k和任意平方n×n矩阵a∈mn（c），我们得到

ρ（a）≤kak.

证据。设λ为λ为最大值的a的特征值，即λ=ρ（a）。如果u（=0）6是与λ相关的任何特征向量，如果u是n×n矩阵，其列均为u，那么au=λu意味着

Au=λu，

从那以后

|λkuk=kλuk=kauk≤kakkuk

而u=06，我们有kuk 6=0，得到

ρ（a）=λ≤kak，

如要求。

命题8.6也适用于mn（r）上的任何实矩阵范数k k，但证明更为微妙，需要诱导范数的概念。在给出定义8.7之后，我们证明了这一点。

结果表明，如果a是实n×n对称矩阵，那么a的特征值都是实的，并且有一些正交矩阵q

a=qdiag（λ1，…，λn）q>，

其中diag（λ1，…，λn）表示其唯一非零项（如果有）是其对角项的矩阵，这是a的（实）特征值。同样，如果a是一个复杂的n×n厄米特矩阵，则a的特征值都是实的，并且存在一些单位矩阵u，因此

a=udiag（λ1，…，λn）u，

其中diag（λ1，…，λn）表示其唯一非零项（如有）是其对角线项的矩阵，这是a的（实）特征值。这些结果的证明见第16章。

现在我们回到矩阵规范。我们从所谓的frobenius范数开始，它就是cn2上的范数k k2，其中n×n矩阵a被视为将a的行（或列）连接在一起得到的向量。读者应该检查任意n×n复矩阵a=（aij）。

.

定义8.6.定义了frobenius范数k kf，使每平方n×n矩阵a∈mn（c），

.

下面的命题表明，Frobenius范数是一个满足其它优良性质的矩阵范数。

提案8.7.Mn（c）上的frobenius norm k kf满足以下特性：

1. 它是一个矩阵范数，即kabkf≤kakf kbkf，对于所有a，b∈mn（c）。
2. 它是幺正不变的，这意味着对于所有的幺正矩阵u，v，我们有

kakf=kuakf=kav kf=kuav kf。

1. pρ（a\_a）≤kakf≤√npρ（a\_a），对于所有a∈mn（c）。

证据。（1）唯一需要证明的属性是事实kabkf≤kakf kbkf。这源于柯西-施瓦兹不等式：

.

1. 我们有

，

和

.

身份

kakf=夸夫肯德基

从前两个开始。

1. 众所周知，矩阵的迹等于其特征值之和。此外，a a是对称半正定的（这意味着它的特征值是非负的），所以ρ（a a）是a a的最大特征值，并且

ρ（a\_a）≤tr（a a）≤nρ（a\_a）、

它通过取平方根产生（3）。

注：弗罗贝尼乌斯范数又称希尔伯特-施密特范数或舒尔范数。这么多与这么简单的事情有关的著名的名字！

## 8.3附属规范

我们现在给出了另一种使用从属规范获得矩阵规范的方法。首先，我们需要一个命题，证明在有限维空间中，矩阵诱导的线性映射是有界的，因此是连续的。

提案8.8.对于≥0c上的每一个范数k k c a，使得n（或rn），对于每一个矩阵a∈mn（c）（或a∈mn（r）），都有一个实常数。

kauk≤ca kuk，

对于每一个向量u∈cn（如果a是实的，则为u∈rn）。

证据。对于cn（或rn）的每个基（e1，…，en），对于每个向量u=u1e1+·····+unen，我们有

kauk=ku1a（e1）+·····+una（en）k

≤u1 ka（e1）k+·················································

式中，c1=max1≤i≤n kca2（e>i）0k。根据定理8.5，kuk1≤c2 k uk对于所有的uk k是等效的，这意味着k k1和k k1是等效的，因此

有一些常数

kauk≤ca kuk，

其中ca=c1c2。

命题8.8表示有限维空间上的每一个线性映射都是有界的。这意味着有限维空间上的每一个线性映射都是连续的。实际上，不难证明赋范向量空间e上的线性映射是有界的，只要它是连续的，不管e的维数是多少。

命题8.8意味着对于每一个矩阵a∈mn（c）（或a∈mn（r）），

.

因为kλuk=λkuk，对于每个非零向量x，我们有

，

这意味着

.

类似地

.

上述考虑证明了以下定义的合理性。

定义8.7.如果k k是cn上的任何范数，我们定义mn（c）上的k kop函数

.

通过范数k.a 7→kakop被称为次矩阵范数或算子范数

矩阵A的算符范数的另一个表示法（特别是Horn和Johnson[92]使用的）是A。

很容易检查函数a 7→kakop是否确实是一个规范，并且根据定义，它满足属性

Kaxk≤Kakop Kxk，对于所有x∈cn。

标准K konocon-mn。由于上述不等式的结果，我们假设满足上述性质的（c）服从于向量。

标准K K K

Kabxk≤Kakop Kbxk≤Kakop Kbxk，

对于所有x∈cn，这意味着

Kabkop≤所有a，b∈mn（c）的Kakop Kbkop，

表明7→Kakop是一个矩阵范数（它是次乘法）。

注意，操作员规范也由

kakop=infλ∈r kaxk≤λkxk，对于所有x∈cn。

c由于函数kcxn−使得yk）和单位spherekxxx=1→k7和xk kaxk=kakop。X

等价地，有一些x∈cn，使得x 6=0和

Kaxk=Kakop Kxk。

运算符规范的定义也意味着

kikop=1.

上述结果表明，Frobenius范数不是一个从属矩阵范数（为什么？）.

如果k k是cn上的向量范数，则它所诱导的算子范数k kop适用于mn（c）中的矩阵。如果我们小心地表示向量和矩阵，这样就不会产生混淆，例如，通过对向量使用小写字母，对矩阵使用大写字母，应该很清楚，kakop是矩阵A的运算符范数，kxk是x的向量范数。因此，遵循MMON练习减轻符号，我们将去掉下标“op”，只写kak而不是kakop。

从属范数的概念可以略作概括。

定义8.8.k k k如果k=r或k=c，对于任意范数k kison m次坐标，n（k），对于任意两个范数k k a和k kb on km，我们说，范数k k k on kn和k kb kaxkb≤kakkxka对于所有a∈mm，n（k）和所有x∈kn。

如果

备注：对于cn上的任何范数k k，我们可以通过以下公式定义mn（r）上的k kr函数：

.

函数a 7→kak r是mn（r）上的矩阵范数，kakr≤kak，

对于所有实矩阵a∈mn（r）。然而，在Cn和实矩阵A上构造向量范数k k是可能的。

Kakr<Kak.

为了避免这类困难，我们在mn（c）上定义了次矩阵规范。幸运的是，对于向量范数k k1、k k2和k k∞，结果是kakr=kak。

我们现在证明命题8.6为实矩阵规范。

提案8.9.对于mn（r）上的任意矩阵范数k k k和任意平方n×n矩阵

a∈mn（r），我们有ρ（a）≤kak。

证据。我们遵循丹尼斯·瑟尔的书[151]中的证据。如果a是实矩阵，问题是与最大模特征值相关的特征向量可能是复杂的。我们使用的技巧基于这样一个事实：对于每个矩阵A（真实或复杂），

ρ（a k）=（ρ（a））k，

剩下的是一个练习（使用命题14.7，它表明如果（λ1，…，λn）是a的（不一定是不同的）特征值，那么（）是k的特征值，对于k≥1）。

在cn上选取任意复杂矩阵范数k kc（例如，frobenius范数，或由cn上的范数诱导的任何从属矩阵范数）。K-k c对实矩阵的约束是一个实范数，我们也用维数n2表示，这里有一个常数c>K-k0c，根据定理8.5，因为N（r）是有限的。

Kbkc≤c Kbk，对于所有b∈mn（r）。

而且，对于每一个，因为k kk≥1，对于每一个实n×n矩阵，我们有一个，根据命题8.6，ρ（ak）≤

是矩阵范数，

，

对于所有k≥1。接下来是

ρ（a）≤c1/k kak，对于所有k≥1。

但是，由于c>0，我们得到limk7→∞c1/k=1（我们得到lim=0）。因此，我们得出结论：

ρ（a）≤kak，

根据需要。

我们现在明确地确定与向量规范k k1、k k2和k∞相关的从属矩阵规范是什么。

提案8.10.对于每个平方矩阵a=（aij）∈mn（c），我们有

.

注意，kak1是a列的1-范数的最大值，k a k是a行的1-范数的最大值。此外，ka k2=kak2，范数∞k k2是单位不变的，这意味着

Kak2=Kuav k2

对于所有的单位矩阵u，v，如果a是一个正规矩阵，那么kak2=ρ（a）。

证据。对于每个向量u，我们有

，

这意味着

.

它仍然表明平等是可以实现的。为此，让j0成为这样一个索引：

，

对于所有i=6 j0和uj0=1，设ui=0。

以同样的方式，我们

，

这意味着

.

为了实现相等，让I0是这样一个索引：

最大x aij=x ai0j。我

J J

读者应该检查



作品。

我们有

kak22=supn kaxk22=supn x a ax.x∈c x∈c

x x=1 x x=1

因为矩阵A A是对称的，所以它具有实特征值，并且可以相对于一个单位矩阵对角化。这些事实可以用来证明函数x 7→x a ax在球面x x=1上的最大值等于a a的最大特征值，即ρ（a a）。我们把证明推迟到讨论优化二次函数为止。因此，

kak2=pρ（a\_a）。

现在使用证明ρ（a a）=ρ（aa）。首先假设ρ（a\_a）>0。在这种情况下，有一些特征向量u（=0）6这样

a a u=ρ（a a）u，

既然ρ（a a）>0，我们必须得到au=06。因为au=06，

a a（au）=a（a au）=ρ（a a）au

这意味着ρ（a\_a）是aa\_的特征值，因此

ρ（a\_a）≤ρ（aa）。

因为（a）=a，用a替换a，我们得到

ρ（a a）≤ρ（a a），

所以ρ（a a）=ρ（aa）。

如果ρ（a a）=0，那么我们必须将ρ（aa）=0，因为根据先前的推理，否则我们将得到ρ（a a）=ρ（aa）>0。因此，无论如何

.

对于任何单位矩阵u和v，证明v a av和a a具有相同的特征值是一个简单的练习，因此

，

以及

.

最后，如果a是一个正规矩阵（aa=a a），则可以证明存在一些单位矩阵u，因此

A=Udu，

其中d=diag（λ1，…，λn）是由a的特征值组成的对角矩阵，因此

a a=（u du）udu=ud u udu=ud du。

然而，d d=diag（λ1 2，…，λn 2），这证明

ρ（a\_a）=ρ（d\_d）=maxλi 2=（ρ（a））2，

我

所以kak2=ρ（a）。

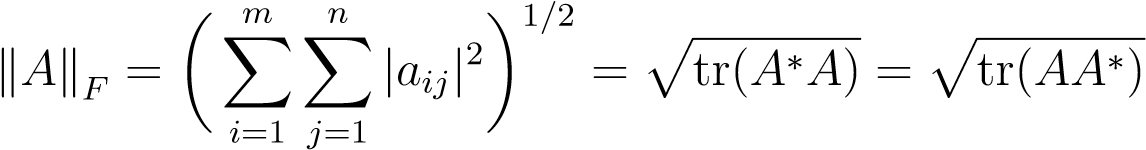
定义8.9.对于a=（aij）∈mn（c），范数kak2=通常称为谱范数。

观察8.7号提案的性质（3）表明

kak2≤kakf≤√nkak2，

这表明，弗罗贝尼乌斯范数是谱范数的上界。弗罗贝尼乌斯范数比谱范数更容易计算。

读者将检查上述证明是否仍然成立，如果矩阵A是真的（将幺正变换为正交），确认向量normsrectangular m×nk kmatries，1，k k2和k k∞的kakr=kak的事实。用同样的公式也很容易验证这个证明是否成立。同样地，由



也是矩形矩阵的范数。对于这些规范，只要AB有意义，我们就得到kabk≤kakkbk。

注：可以看出，对于任意两个实数p，q≥1，当=1时，我们得到ka kq=kakp=sup<（y ax）kxkp=1，kykq=1 sup hax，yi kxkp=1，kykq=1，其中ka kq和kakp是操作规范。

注：设（e，k k）和（对于范数onf，k）为两个赋范向量空间（为便于表示，e和f，这不应引起任何混淆）。我们用同样的符号来回忆，函数f k k:e→f是连续的，如果对于每一个a∈e，对于每一个>0，有一些η>0，这样对于所有x∈e，

如果kx−ak≤η，则

不难证明线性图f:e→f是连续的，如果有常数c≥0，那么

kf（x）k≤c kxk，对于所有x∈e。

如果是这样的话，我们就说从f开始的所有连续（等价的，有界）线性映射的集合是有界的（或线性有界算子）。我们删除到f，然后我们可以定义（e；f）表示

l（e；f）上的算符范数（或次范数）k k如下：对于每个f∈l（e；f），

，

或等同于

kf k=infλ∈r kf（x）k≤λkxk，对于所有x∈e。

不难看出图F7→Kfk是满足该特性的L（e；f）上的一个标准。

kf（x）k≤kfkkkxk

对于所有x∈e，如果f∈l（e；f）和g∈l（f；g），则kg fk≤kgkkkkfk。

算子规范在函数分析中起着重要作用，尤其是当空间e和f是完整的时。

## 8.4涉及从属规范的不平等

在本节中，我们将讨论本章最后三节中某些证明所需的两个技术不等式。首先，我们证明了当我们处理矩阵的条件数时需要的一个命题。

提案8.11.设k k为任意矩阵范数，设b∈mn（c），使kbk<1。（1）如果k k是次矩阵范数，那么矩阵i+b是可逆的，并且

.

（2）如果I+B形式的矩阵是奇异的，那么对于每个矩阵范数（不一定是从属的），Kbk≥1。

证据。（1）观察（i+b）u=0表示bu=-u，所以kuk=kbuk。

回想一下

Kbuk≤Kbkkuk

为每一个下属规范。因为kBk<1，如果u 6=0，那么kBk<kuk，

这与kuk=kbuk相矛盾。因此，我们必须有u=0，证明i+b是内射的，因此是双射的，即可逆的。然后我们有了

（i+b）−1+b（i+b）−1=（i+b）（i+b）−1=i，

8.4。涉及从属规范的不等式

会产生

，

最后，

.

（2）如果i+b是奇异的，那么−1是b的特征值，根据命题8.6，我们得到ρ（b）≤kbk，这意味着1≤ρ（b）≤kbk。

第二个不等式是处理矩阵幂序列收敛性所需要的结果。

提案8.12。对于每一个矩阵a∈mn（c），对于每一个，都有一些从属矩阵范数k k，这样



证据。根据定理14.5，存在一些可逆矩阵u和一些上三角矩阵t，从而

A=UTU−1，

说吧。

λ1 t12 t13··t1n

0λ2 t23···t2n

T=………………，

γ

0 0···λn−1 tn−1n

0 0···0λn

其中，λ1，…，λn是a的特征值。对于每个δ=06，定义对角矩阵

dδ=diag（1，δ，δ2，…，δn-1）

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

现在定义函数k k:mn（c）→r by

，

每b∈mn（c）。然后，很容易验证上述函数是从属于向量范数的矩阵范数

.

此外，对于每>0，我们可以选择δ，以便

，

根据范数k k的定义，我们得到∞



这表明我们所构造的范数满足所需的属性。注意等式通常是不可能的；考虑矩阵

，

其中，ρ（a）=0<kak，因为a=0.6

## 8.5矩阵的条件数

不幸的是，存在线性系统ax=b，其解在b或a的小扰动下不稳定。

.

读者应该检查它是否有解决方案x=（1,1,1,1）。如果我们稍微干扰右边的b+∆b，其中

，

我们得到了新的系统

.

新的解决方案是x+∆x=（9.2、-12.6、4.5、-1.1），其中

∆X=（9.2、−12.6,4.5、−1.1）−（1,1,1,1）=（8.2、−13.6,3.5、−2.1）。然后是数据相对于一个范数的相对误差，

，

在输入中产生相对错误

.

因此，数据中1/300阶的相对阶数在解中产生7/1阶的相对误差，这表示2100阶相对误差的放大。现在让我们稍微扰动矩阵，得到新的系统。

.

这一次，解决方案是x+∆x=（−81137、−34,22）。同样，数据中的一个小变化会极大地改变结果。然而，原始系统是对称的，有行列式1，并且有整数项。问题是系统的矩阵条件很差，我们现在将解释这个概念。

给定一个可逆矩阵A，首先假设我们将b扰动到b+∆b，然后让我们分析两个系统的两个精确解x和x+∆x之间的变化。

AX＝B

A（X+∆X）=B+∆B。

我们还假设我们有一些范数k k，并且我们在矩阵上使用从属矩阵范数。从

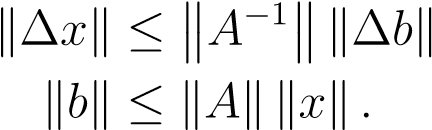
AX＝B

a x+a∆x=b+∆b，

我们得到

∆x=A−1∆b，

我们得出结论



因此，结果k∆xk/kxk中的相对误差以数据中的相对误差k∆bk/kbk为界，如下所示：

.

现在让我们假设A受A+的扰动，让我们分析两个系统的精确解之间的变化。

AX＝B

（a+∆a）（x+∆x）=b.

第二个方程得出a x+a∆x+∆a（x+∆x）=b，减去第一个方程，我们得到

接下来是

，

可以重写为

.

请注意，即使矩阵a+∆a是奇异的，上述推理也是有效的，只要不合理地期望Ratiox+∆x是第二个系统的解。此外，如果k∆xk/kx+∆xk足够小，那么它就是∆xk/kxk。稍后会更精确地说明这一点。

总之，对于这两个扰动中的每一个，我们发现结果因子中的相对误差是最优的，这表明了以下定义：以数据中的相对误差为界，乘以Kakka-1K。事实上，这

定义8.10.对于任何从属矩阵范数k k，对于任何可逆矩阵a，数字condition（

称为相对于k k的条件数。

条件数cond（a）测量线性系统ax=b对数据b和a变化的敏感性；一种被称为系统条件的特征。因此，当我们说一个线性系统是病态的时，我们的意思是它的矩阵的条件数很大。我们可以将前面的分析尖锐化如下：

提案8.13.设a为可逆矩阵，设x和x+为线性系统的解。

AX＝B

A（X+∆X）=B+∆B。

如果b 6=0，那么不等式

康德

是最好的。这意味着对于给定的矩阵A，存在一些向量b 6=0和∆b6=0，其中等式成立。

有一些向量机存在。我们已经证明了不平等。现在，因为x 6=0和∆b 6=0，其中k k是次矩阵范数，

和kaxk=kakkxk。

提案8.14.设a为可逆矩阵，设x和x+为两个系统的解。

AX＝B

（a+∆a）（x+∆x）=b.

如果b 6=0，那么不等式

康德

是最好的。这意味着在给定矩阵A的情况下，存在一个向量b 6=0和一个矩阵立场，如果k∆∆aak=06<1/等式成立。此外，如果kA−1K），我们得到k∆ak足够小（对于

条件（））；

事实上，我们有康德。

证据。第一个不平等已经被证明。为了证明可以实现相等，让w是任何向量，使得w 6=0和

，

设β6=0为任意实数。现在向量

∆x=−βa−1w x+∆x=w

B=（A+βi）W

和矩阵

∆a=βi

使方程变得时髦

.

最后，我们可以选取β，这样它就不等于a的任何特征值，所以a+∆a=a+βi是可逆的，并且是非零的。

如果k∆ak<1/ka−1K，则

，

因此，根据命题8.11，矩阵i+a−1∆a是可逆的，并且

.

回想一下我们之前证明的

∆x=−a−1∆a（x+∆x）、

把x加到两边，把右边移到左边，得到

（i+a−1∆a）（x+∆x）=x，

因此

X+∆X=（I+A−1∆A）−1X，

会产生

∆x=（i+a−1∆a）−1−i）x=（i+a−1∆a）−1（i−（i+a−1∆a））x=−（i+a−1∆a）−1a−1（∆a）x。

从这个和

，

我们得到

可以写为

康德

这就是我们所寻求的那种不平等。

注：如果A和B同时受到扰动，则得到“扰动”系统。

（a+∆a）（x+∆x）=b+∆b，

可以看出，如果k∆ak<1/ka−1K（且b=0）6，那么

；

见德梅尔[49]第2.2节和霍恩和约翰逊[92]第5.8节。

我们现在列出条件数的一些性质，并计算出光谱范数（k k2引起的矩阵范数）的cond（a）是什么。首先，我们需要引入一个非常重要的矩阵因子分解，即奇异值分解，简而言之，SVD。

可以看出（见第20.2节），在任意n×n矩阵a∈mn（c）下，存在两个单位矩阵u和v，一个实对角矩阵∑=diag（σ1，…，σn），其中σ1≥σ2≥······························

A=V∑U。

定义8.11.给定一个复n×n矩阵a，使a=v∑u>的三重（u，v，∑）式，其中u和v是n×n的一元矩阵，∑diag（σ1，…，σn）是实数的对角矩阵，σ1≥σ2≥········································u和v是正交矩阵，非负数σ1，…，σn被称为a的奇异值。

因式分解a=v∑u意味着

A A=U∑2U和AA=V∑2V，

这表明A A和AA的特征值，U列是A A对应的eivenvectors，V列是AA对应的eivenvectors。

由于是a a（和aa）的最大特征值，请注意pρ（a a）=pρ（aa）=σ1。

推论8.15。矩阵A的谱范数kak2等于a的最大奇异值。同样，矩阵A的谱范数kak2等于矩阵A的∞范数。

奇异值向量，

.

因为矩阵a的frobinius范数由kakf=ptr（a a）定义，并且

TR（

A A的特征值在哪里，我们看到了

.

推论8.16。矩阵的Frobenius范数由其奇异值向量的2-范数给出；kakf=k（σ1，…，σn）k2。

对于正态矩阵，如果λ1，…，则λn是a的（复杂）特征值，则

σi=λi，1≤i≤n。

提案8.17。对于每一个可逆矩阵a∈mn（c），下列性质成立：

（1）

cond（a）≥1，cond（a）=cond（a−1）cond（αa）=cond（a）表示所有α∈c−0。

（2）如果cond2（a）表示a相对于光谱范数的条件数，那么

康德

式中，σ1≥···········≥σn为a（3）的奇异值，如果矩阵a为正态，则

康德

式中，λ1，…，λn是a的特征值，因此λ1≥····≥λn。

1. 如果a是一元矩阵或正交矩阵，则

条件2（a）=1.

1. 条件数cond2（a）在幺正变换下是不变的，这意味着cond2（a）=cond2（ua）=cond2（av）。

对于所有的幺正矩阵u和v。

证据。（1）中的性质是从属矩阵范数性质的直接后果。特别是，aa−1=i意味着

=康德（A）。

1. 我们之前已经证明了），这是a a最大特征值的模的平方。因为我们刚刚看到，其中，σ1，…，σn的特征值是a的奇异值，我们有

Kak2=σ1。

如果a是可逆的，那么σ1≥·····················································

，

因此

康德

1. 这是因为对于正态矩阵，kak2=ρ（a）。
2. 如果a是一个单位矩阵，那么a a=aa=i，那么ρ（a a）=1，和kak2=

pρ（a\_a）=1.我们还有kA−1k2=kA k2=pρ（a a）=1，因此cond（a）=1。

1. 这直接来自光谱范数的一元化不变性。

命题8.17（4）表明，幺正变换和正交变换是非常好的条件，第（5）部分表明，幺正变换保留了条件数。

为了计算cond2（a），我们需要计算a的顶部和底部奇异值，这可能很困难。不等式kak2≤kakf≤√nkak2，

如果可以确定a−1，则在获得cond2（a）=kak2 ka−1k2的近似值时可能很有用。

备注：Cond2（a）有一个有趣的几何特征。如果θ（a）表示所有正交向量对上向量a u和a v之间的最小角度，作为u和v范围，则可以证明

cond2（a）=cot（θ（a）/2））。

因此，如果a接近奇异，那么会有一些正交对u，v，使得au和av接近平行；角度θ（a）将是小的，cot（θ（a）/2）将是大的。有关更多详细信息，请参见Horn和Johnson[92]（第5.8节和第7.4节）。

应该注意的是，一般情况下（如果a不是一个正规矩阵），一个矩阵可能有一个非常大的条件数，即使它的所有特征值都是相同的！例如，如果我们考虑n×n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 |

结果表明，条件2（a）≥2N−1。

具有非常大条件数的矩阵的经典例子是希尔伯特矩阵H（n），n×n矩阵

例如，当n=5时，

.

可以看出，cond2（h（5））约为4.77×105。

希尔伯特于1894年在研究近似理论中的一个问题时引入了这些矩阵。希尔伯特矩阵h（n）是对称正定的。可以给出其行列式的闭式公式（它是所谓的柯西行列式的一种特殊形式）；见问题8.15。h（n）的倒数也可以明确计算；见问题8.15。可以看出

Cond2（H（N））=O（（1+√2）4N/√N）。

回到我们的矩阵

，

它是一个对称正定矩阵，可以证明它的特征值，在这种情况下，由于a是spd，它也是它的奇异值，是

λ1≈30.2887>λ2≈3.858>λ3≈0.8431>λ4≈0.01015，

所以这个条件。

读者应检查，对于之前使用的右侧b的扰动，相对误差k∆xk/kxk和k∆xk/kxk满足不等式。

康德

接近平等。

8.6。规范的应用：不一致线性系统

## 8.6规范的应用：解决不一致线性系统

求解不一致线性系统ax=b的问题在实践中经常出现。这是一个b不属于a列空间的系统，通常方程多于变量。因此，这样的系统没有解决方案。然而，我们仍然希望“解决”这样一个系统，至少大致上是这样。

这样的系统通常在试图适应某些数据时出现。例如，我们可能有一组三维数据点

P1，…，PN，

我们有理由相信这些点几乎是共面的。我们想找到一个最适合我们数据点的平面。回想一下，平面方程是

αx+βy+γz+δ=0，

（α，β，γ）=（06，0,0）。因此，每个平面要么不平行于x轴（α=0）6，要么不平行于y轴（β=0）6，要么不平行于z轴（γ=0）。

假设我们有理由相信我们要找的平面不平行于z轴。如果我们错了，在最小二乘解中，其中一个系数，α，β，将会非常大。如果γ=06，那么我们可以假设我们的平面是由一个形式方程给出的。

Z=ax+x+d，

我们希望这个方程满足所有的π，这导致n个方程组在3个未知数a，b，d中，具有π=（Xi，Yi，Zi）；

ax1+by1+d=z1……

axn+byn+d=zn。

然而，如果n大于3，这样的系统通常没有解决方案。由于上述系统不能完全解决，我们可以尝试找到一个解决方案（A，B，D），使

最小二乘误差

.

这就是勒让德和高斯在19世纪初发现的！

一般来说，给定一个线性系统

ax=b，

我们解决了最小二乘问题：最小化。

幸运的是，每个n×m矩阵a都可以写成

A=V度>

其中u和v是正交的，d是具有非负项的矩形对角矩阵（奇异值分解或SVD）；见第20章。

SVD可以用来解决不一致的系统。如第21章所示，存在一个最小范数的向量x，使kax−bk2最小化。它由（penrose）的伪逆（本身由SVD给出）给出。

据观察，在最小二乘意义上的求解可能会给“异常值”带来过多的权重，也就是说，在最佳拟合平面之外的点。在这种情况下，最好是最小化（1-范数）

.

这似乎不是线性问题，但我们可以使用一个技巧将这个最小化问题转换为线性程序（这意味着一个涉及线性约束的问题）。

注意x=max x、−x。因此，通过引入新的变量e1，…，en，我们的最小化问题等价于线性规划（lp）：

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

注意约束条件等于

ei≥axi+byi+d−zi，1≤i≤n。

对于一个最优解，我们必须有等式，否则我们可以减少一些ei，得到一个更好的解。当然，我们不再处理“纯”线性代数，因为我们的约束是不等式。

我们现在不喜欢学习线性规划，但是上面的例子提供了一个学习更多线性规划的好理由！

## 8.7序列和序列的限制

如果x∈r或x∈c和if/（1x−<x1），众所周知，当n趋于无穷大时，我们写

收敛到极限1

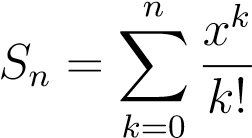
.

8.7。序列和序列的极限

例如，

.

同样地，总和



当n变为无穷大时，每x（r或c）收敛到ex。如果我们用复数n×n矩阵a的实数替换x呢？

部分和仍然有意义，但我们必须定义矩阵序列的极限。这可以在任何赋范向量空间中完成。

定义8.12.设（e，kk）为赋范向量空间。e中的序列（u n）n∈n是任意函数u:n→e，对于任意v∈e，序列（un）收敛到v（v是序列（un）的极限），如果对于每>0，有一个整数n>0，这样

对于所有n≥n。

通常我们假设一个序列被n−0索引，也就是说，它的第一个项是u1而不是u0。

如果序列（un）收敛到v，那么由于三角形不等式

kum−unk≤kum−vk+kv−unk，

我们可以看到，对于每>0，我们可以找到n>0，这样2，所以

对于所有m，n≥n。

上述性质对于收敛序列是必要的，但不一定足够。例如，如果e=q，有满足上述条件的理性序列，但其极限不是有理数。例如，序列收敛到e，序列收敛到π/4，但e和π/4不是理性的（事实上，它们是超越的）。然而，R是由Q构造的，以保证具有上述性质的序列收敛，C也是。

定义8.13.给定一个赋范向量空间（e，k k），序列（un）是一个柯西序列，如果对于每>0，有一些n>0，这样

对于所有m，n≥n。

如果每个柯西序列都收敛，那么我们就说e是完整的。完全赋范向量空间也称为Banach空间。

R的一个基本性质是它是完整的。紧接着，C也完成了。如果e是一个有限维实向量空间或复向量空间，由于任意两个范数相等，我们可以选取∞范数，然后通过选取e中的一个基，e中的一个向量序列（un）收敛，如果n个坐标序列（）收敛，那么任何有限维实向量或复向量空间E是一个巴拿赫空间。

现在我们来考虑级数的收敛性。

定义8.14.给定一个赋范向量空间（e，k k），一个级数是元素的无穷和Uk∈e，我们用sn表示第一个n+1元素的部分和，

.

定义8.15.我们说，如果序列（sn）收敛到v，也就是说，给定任何>0，存在一个正整数n，这样对于所有的

n≥n，



在这个例子中，我们说这个级数是收敛的。我们说，如果一系列规范是收敛的，那么这个系列绝对收敛。

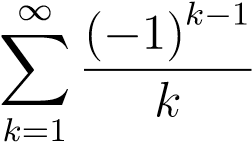
如果级数收敛到v，因为对于所有m，n和m>n，我们有

，

如果我们让m去无穷大（n固定），我们会看到序列号，和

.

有些序列是收敛的，但不是绝对收敛的；例如，序列



收敛到ln2，但不收敛（这个和是无限的）。

如果e是完整的，那么相反的结果是非常有用的。

提案8.18。假设（e，k k）是一个完全赋范向量空间。如果一个级数绝对收敛，那么它就是收敛的。

8.8。矩阵指数

证据。如果绝对收敛，则证明该序列（sm）是一个柯西序列，也就是说，对于每个>0，有一些p>0，这样对于所有n≥m≥p，



注意

ksn−smk=kum+1+····+unk≤kum+1k+····+kunk，

由于序列收敛，它满足柯西准则。因此，序列（sm）也满足柯西准则，并且由于e是一个完整的向量空间，因此序列（sm）收敛。

注：可以看出，如果（e，k k）是一个赋范向量空间，使得每个绝对收敛级数也收敛，那么e必须是完整的（见Schwartz[146]）。

绝对收敛的一个重要推论是，如果数列中的项被重新排列，那么所得的数列仍然绝对收敛并且具有相同的和。更准确地说，让σ是自然数的任意置换（双射）。这个序列称为原始序列的重新排列。可以显示以下结果（见Schwartz[146]）。

提案8.19。假设（e，k k）是赋范向量空间。如果一个级数是收敛的也是绝对收敛的，那么对于n的每一个置换σ，该级数都是收敛的和绝对收敛的，其和等于原始级数的和：

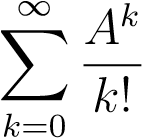
.

特别地，如果（e，k k）是一个完全赋范向量空间，那么命题8.19成立。

现在我们将8.18号命题应用于矩阵指数。

## 8.8矩阵指数

提案8.20。对于任意n×n实矩阵或复矩阵a，序列



在mn（c）（或mn（r））上绝对收敛于任何算子范数。

证据。选择cn（或rn）上的任何范数，并让kk为mn（c）上的相应运算符范数。因为mn（c）的尺寸为n2，所以它是完整的。通过命题8.18，它足以证明非负实级数收敛。因为k k是一个算子范数，这是一个矩阵范数，所以我们有

.

因此，正实数的非递减序列以ekak为界，并以r的一个基本性质为界，它有一个最小上界，即它的极限。

定义8.16.设e为复赋范向量空间的有限维实。对于任意n×n矩阵a，序列的极限



是a的指数，表示为ea。

指数x 7→x的一个基本性质是

ex+y=exey，对于所有x，y∈c。

因此，e x总是可逆的，（ex）−1=e−x。对于矩阵，因为矩阵乘法一般不可交换，

ea+b=eaeb

失败！这个结果被挽救如下。

提案8.21。对于任意两个n×n复矩阵a和b，如果a和b上下班，即ab=ba，那么ea+b=eaeb。

8.21号提案的证明见Gallier[73]。

由于a和−a通勤，作为命题8.21的推论，我们看到ea总是可逆的，并且

（e a）−1=E−A。

也很容易看出

（ea）>=ea>。

8.8。矩阵指数

一般来说，矩阵A的指数ea没有闭式公式，但对于维2和维3的斜对称矩阵，有显式公式。每个人都应该喜欢计算指数ea

如果我们写信

，

然后

关键属性是

J2=−I.

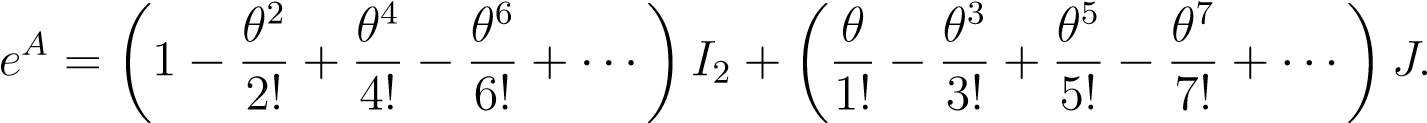
提案8.22。如果a=θj，则

.

证据。我们有，所以

.

我们重新安排了条款的顺序



我们识别cosθ和sinθ的幂级数，因此

那就是

，

如要求。

因此，我们发现2×2的斜对称矩阵的指数是旋转矩阵。此属性可归纳为任何维度。第11.7节给出了n=3（罗德里格斯公式）时的显式公式。

提案8.23。如果b是n×n（实）斜对称矩阵，即b>=−b，则

q=eb是一个正交矩阵，即

Q>Q=QQ>=I。

证据。既然b>=-b，我们有

q>=（e b）>=eb>=e−b。

自从B和B通勤后，我们

q>q=e−b e b=e−b+b=e0=i。

同样地，

qq>=ebe−b=eb−b=e0=i，

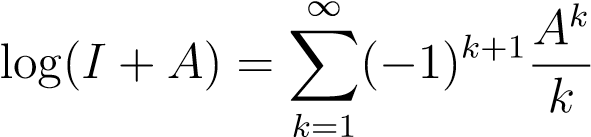
这就是证据的结论。

也可以证明，det（q）=det（eb）=1，但这需要更好地理解eb的特征值（见第14.5节）。此外，对于每个n×n旋转矩阵q（正交矩阵q，使得det（q）=1），都有一个斜对称矩阵b，使得q=eb。这是一个基本性质，在机器人学中有n=3的应用。

所有熟悉的系列都有类似的矩阵。例如，如果k a k<1（其中k k是一个算符范数），那么序列绝对收敛，可以证明它的极限是（i−a）−1。

另一个有趣的系列是对数。对于任意n×n复矩阵a，如果kak<1

（其中k k是一个算子范数），然后是序列



绝对收敛。

## 8.9总结

本章的主要概念和结果如下：

* 规范和赋范向量空间。
* 三角形不等式。

8.9。总结

* 欧几里得准则；p-准则。
* H–Older的不平等；Cauchy–Schwarz的不平等；Minkowski的不平等。
* 厄米田内积和欧几里得内积。
* 等效规范。
* 有限维向量空间上的所有范数都是等价的（定理8.5）。
* 矩阵规范。
* 赫米特矩阵、对称矩阵和正规矩阵。正交矩阵和幺正矩阵。
* 矩阵的轨迹。
* 矩阵的特征值和特征向量。
* 矩阵的特征多项式。
* 矩阵A的光谱半径ρ（a）。
* 弗罗贝尼乌斯准则。
* 弗罗贝尼乌斯范数是一个统一不变的矩阵范数。
* 有界线性映射。
* 从属矩阵规范。
* k k向量范数的次矩阵范数的特征∞.k k1、k k2和
* 光谱标准。
* 对于每一个矩阵k，如果a∈mn（c），对于每一个大于0的矩阵，都有一些次矩阵。
* 矩阵的条件数。
* 线性系统的摄动分析。
* 奇异值分解（SVD）。
* 条件编号的属性。A.2（a）的最大和最小奇异值的特征
* 希尔伯特矩阵：一个非常糟糕的条件矩阵。
* 用最小二乘法求解不一致线性系统；线性规划。
* 赋范向量空间中向量序列的收敛性。
* 柯西序列，复赋范向量空间，Banach空间。
* 级数收敛。绝对收敛。
* 矩阵指数。
* 斜对称矩阵和正交矩阵。

## 8.10问题

问题8.1。设A为下列矩阵：

.

计算a的算符2-范数kak2。

问题8.2。证明命题8.3，即下列不等式对所有x∈rn（或x∈cn）都成立：

，

问题8.3。对于任意p≥1，证明对于所有x∈rn，

plim kxkp=kxk∞。

→∞7

问题8.4.设A为严格对角占优的n×n矩阵，即

对于i=1，…，n，和let

.

严格的行对角占优的事实等于条件δ>0。（1）对于任何非零矢量v，证明

kVk∞≥kVk∞δ。

用上面的例子来证明a是可逆的。

（2）证明

暗示。证明这一点

.

问题8.5。设A为任意可逆复数n×n矩阵。

1. 对于Cn上的任何向量范数k k，证明由

kxka=所有x∈cn的kxk，

是向量范数。

1. 证明了由k ka（也用k ka表示）引起的算子范数由下式给出：

对于每个n×n矩阵b，

其中kaba−1K使用由k k引起的算符范数。

问题8.6.给出一个关于cn和实矩阵a的范数的例子

Kakr<Kak，

其中，k−kr和k−k是与向量范数k−k相关的运算符范数。

暗示。这可以在n=2时完成。

问题8.7.c进一步证明，如果=1/（2k a−1k）let，那么对于everyk k，k是任意的算符范数。k n×n矩阵（a+hh）给出可逆的，如果−1k≤kh1k≤/c.c，那么an+×hn矩阵是可逆的。a，如果

hk≤c，然后k

问题8.8.设a为任意m×n矩阵，设λ∈r为任意正实数λ>0。

1. 证明a>a+λin和aa>λim是可逆的。
2. 证明这一点

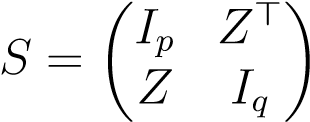
A>（a a>+λim）−1=（a>a+λin）−1a>。

注：上述表达式与函数所对应的矩阵相对应。

Φ（x）=（ax−b）>（ax−b）+λx>x

达到最小值。它出现在机器学习（内核方法）中。

问题8.9.设z为q×p实矩阵。证明如果ip−z>z是正定的，那么（p+q）×（p+q）矩阵



是对称正定的。

问题8.10。证明对于任何实矩阵或复矩阵A，我们有

，

其中，上述规范为运营商规范。

暗示。使用命题8.10（除其他外，它表明

问题8.11。说明图A 7→ρ（a）（其中，ρ（a）是a的光谱半径）既不是范数，也不是矩阵范数。特别是，找到两个2×2矩阵a和b，这样

ρ（a+b）>ρ（a）+ρ（b）=0且ρ（ab）>ρ（a）ρ（b）=0。

问题8.12。定义图a 7→m（a）（在n×n实矩阵或复杂n×n矩阵上定义）的方法是

m（a）=max a i j 1≤i，j≤n。

1. 证明这一点

m（a b）≤nm（a）m（b）

对于所有n×n矩阵a和b。

1. 给出一个不等式的反例

m（a b）≤m（a）m（b）。

1. 证明地图A 7→Kakm由

kakm=nm（a）=nmax a i j 1≤i，j≤n

是矩阵范数。

问题8.13。设为实对称正定矩阵。

1. 利用Cholesky因式分解证明了存在一些上三角矩阵c，如果其对角元素严格为正，则它是唯一的，例如s=c>c。
2. 对于任何x∈rn，定义kxk=（x>sx）1/2。

S

证明kxks=kcxk2，

地图x 7→kxks是一个标准。

问题8.14。设A为实数2×2矩阵

.

1. 证明A的奇异值σ1≥σ2的平方是

二次方程

x2−tr（a>a）x+det（a）2=0.

1. 如果我们让

，

证明条件。

1. 考虑2×2可逆矩阵的子集，其条目aij是整数，因此0≤aij≤100。

证明在相同的a值下，函数cond2（a）和礹（a）在集合s上达到最大值。

检查矩阵的那个

我们有

和cond2（am）≈39206。

1. 证明对于所有a∈s，如果det（a）≥2，则（a）≤10000。得出结论，对于矩阵，s上的最大μ（a），使得Det（a）=1。证明求S上最大μ的矩阵等于求一些整数n1、n2、n3、n4，这样

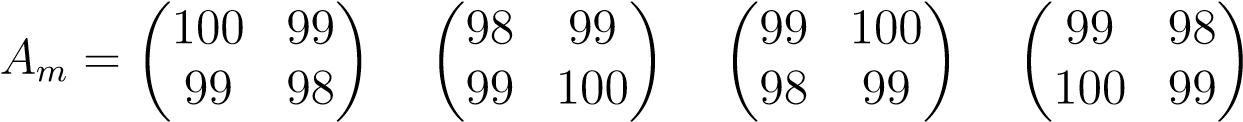
0≤n4≤n3≤n2≤n1≤100

N21+N22+N23+N24≥1002+992+992+982=39206 N1N4−N2N3=1.

您可以在没有证据的情况下使用，事实上，对上述约束的唯一解决方案是多集

\_100,99,99,98\_

1. 从第（4）部分中推断，μ具有最大值的s中的矩阵为



检查这些矩阵的μ值是否相同。得出结论

最大条件2（A）=条件2（AM）。

阿斯

1. 解决系统问题

.

干扰右侧B

并求解新系统，其中y=（y1，y2）。检查那个

.

计算k xk2、k∆xk2、kkk2、k∆bk2，并估算

.

检查c≈cond2（am）=39206。

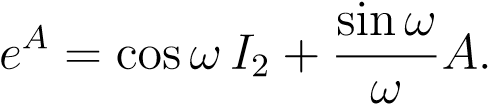
问题8.15。考虑一个实数2×2矩阵，其形式为零。

1. 证明这一点

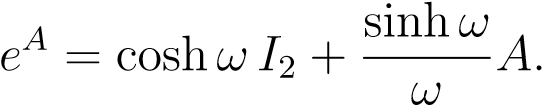
.

如果a2+bc=0，证明ea=i2+a。

1. 如果a2+bc<0，设ω>0，使ω2=−（a2+bc）。证明这一点



1. 如果a2+bc>0，让ω>0等于ω2=a2+bc。证明这一点



1. 证明在所有情况下

=1和Tr（a）≥−2。

1. 证明了存在一些实2×2矩阵b，且det（b）=1，因此没有实2×2矩阵a的零迹，因此ea=b。

问题8.16。还记得希尔伯特矩阵是由

1. 证明这一点

，

整数的倒数。

暗示。使用问题？？

1. 令人惊讶的是，h（n）的倒数项是整数。证明（h（n））-1=（αij），与

.

294第8章。向量范数和矩阵范数

第九章

# 求解线性系统的迭代法

## 9.1向量和矩阵序列的收敛性

在第七章中，我们讨论了求解线性方程组的一些主要方法。这些方法是直接的方法，因为它们产生精确的解（假设无限精确！）.

求解线性系统的另一类方法包括用迭代法近似解。其基本思想是：给定一个线性系统ax=b（在mn（c）中有一个平方可逆矩阵），求另一个矩阵b∈mn（c）和一个向量c∈cn，这样

1。矩阵i-b是可逆的2。系统的唯一解x ax=b与E系统的唯一解u相同。

u=bu+c，

然后从任意向量u0开始计算

UK+1=Buk+C，K∈N。

在某些条件下（即将澄清），序列（UK）收敛到极限Ue，这是u=bu+c的唯一解，因此ax=b。

因此，找到确保上述序列收敛的条件，并利用工具比较这些序列的收敛速度是非常重要的。因此，我们从一些向量和矩阵序列收敛的一般结果开始。

二百九十五

设（e，k）为赋范向量空间。从第8.7节回忆起，向量u k∈e的序列（uk）收敛到极限u∈e，如果对于每个>0，有一些自然数n，这样对于所有k≥n。

我们写信

U=Lim英国。K7→∞

如果e是一个有限维向量空间，dim（e）=n，我们从定理8.5中知道任意两个范数都是等价的，如果我们选择范数k k∞，我们可以看到，向量序列的收敛性uk等于由分量构成的n个标量序列的收敛性。f这些向量（在任何基础上）。同样的性质也适用于m×n矩阵（k=r或k=c）的有限维向量空间mm，n（k），这意味着矩阵序列的收敛性等于m×n标量序列的收敛性（），其中i，j固定（1≤i≤m，1≤j≤n）。

下面的第一个定理给出了矩阵B的幂序列（bk）收敛到零矩阵的一个充要条件。回想一下，矩阵b的光谱半径ρ（b）是b特征值的模λi\_的最大值。

定理9.1。对于任何平方矩阵b，以下条件是等效的：

1. limk7→∞bk=0，
2. limk7→∞bkv=0，对于所有向量v，
3. ρ（b）<1，
4. Kbk<1，对于某些次矩阵范数k k k。

证据。假设（1），k k是e上的向量范数，k是相应的矩阵范数。对于每一个向量v∈e，因为k k是一个矩阵范数，我们有

Kbkvk≤Kbkkkvk，

也就是说，利芒自limk7→∞k7→∞bbkvk=0=0。这证明（1）假设lim k 7→∞kbkk=0，我们得出lim（2）.k→∞7kbkvk=0，

假设（2）。如果ρ（b）≥1，那么会有一些特征向量u（=0）6和一些特征值λ，这样

bu=λu，λ=ρ（b）≥1，

但序列（bku）不会收敛到0，因为bku=λku且λk=λk≥1。由此可见（2）暗示（3）。

假设（3）成立，即ρ（b）<1。通过命题8.12，我们可以找到大于0的足够小的1，以及一个从属矩阵范数k，这样



9.1。向量和矩阵序列的收敛性

即（4）。

最后，假设（4）。因为k k是矩阵范数，

KBkk≤KBkk，

由于Kbk<1，我们推断（1）成立。

研究迭代法的收敛速度需要以下命题。

提案9.2.对于每一个平方矩阵b∈mn（c）和每一个矩阵范数kk，我们有

lim kbkkk1/k=ρ（b）。

K7→∞

证据。我们从命题8.6中知道，ρ（b）≤k b k，并且由于ρ（b）=（ρ（bk））1/k，我们推导出所有k≥1的ρ（b）≤kbkkk，

如此

ρ（b）≤lim kbkkk1/k。

K7→∞

现在让我们证明，对于每个>0，都有一个整数），这样

对所有人来说，

这证明了

lim kbkkk1/k≤ρ（b），

K7→∞

以及我们的建议。

对于任何给定的矩阵

.

因为1，定理9.1意味着lim=0。因此，有一个整数），这样对于所有人来说

，

这意味着，正如所声称的。

我们现在将上述结果应用于迭代方法的收敛性。

## 9.2迭代法的收敛性

回想一下，求解线性系统a x=b（a∈mn（c）可逆）的迭代方法包括找到一些矩阵b和一些向量c，这样i-b是可逆的，ax=b的唯一解x等于u=bu+c的唯一解u，然后从任意向量开始eu0，计算由

UK+1=Buk+C，K∈N，

并且说迭代法是收敛的iff

klim7→∞UK=U，E

对于每个初始向量u0。

这里是基于矩阵B的任何迭代方法收敛的基本准则，称为迭代方法的矩阵。

定理9.3.如果系统u=bu+c如上所述，其中i−b是可逆的，则以下陈述是等效的：

1. 迭代法是收敛的。
2. ρ（b）<1.
3. Kbk<1，对于某些次矩阵范数k k k。证明。通过以下方式定义矢量Ek（误差矢量）

Ek=英国-U，E

其中u是系统的唯一解u=bu+c。显然，迭代法是

e

收敛iff

Lim Ek=0.K7→∞

我们声称

Ek=BKE0，K≥0，

其中e0=u0−ue。

通过对k的归纳证明了这一点。基本情况k=0是微不足道的。根据诱导假设，Ek=bke0，由于Uk+1=buk+c，我们得到

UK+1−Ue=Buk+C−U，E

由于Ue=bue+c，Ek=bke0（通过诱导假设），我们得到

UK+1−Ue=b uk−Bue=b（UK−Ue=bek=bbke0=bk+1e0，

验证导入步骤。因此，迭代法收敛于iff

lim bke0=0.K7→∞

因此，我们的定理遵循定理9.1。

9.2。迭代法的收敛性

需要下一个命题来比较迭代方法的收敛速度。结果表明，误差向量Ek=bke0的渐近行为最差为（ρ（b））k。

提案9.4.设k k为任意向量范数，设b∈mn（c）为i−b可逆的矩阵，设u为u=bu+c.e的唯一解。

1. if（uk）是迭代定义的任何序列

然后

.

1. 假设b1和b2是两个矩阵，使得i−b1和i−b2是可逆的，假设u=b1u+c1和u=b2u+c2都有相同的唯一解ue，并考虑由

UK+1=B1UK+C1

vk+1=b2vk+c2，

U0=v0。如果ρ（b1）<ρ（b2），那么对于任何一个，都有一个整数，这样

对所有人来说，我们有

.

证据。设k k为次矩阵范数。回想一下

UK−Ue=黑色0，

当e0=u0-ue时。对于每一个k∈n，我们有

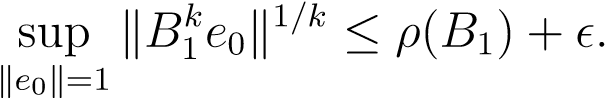
，

这意味着

，

陈述（1）来自9.2号提案。因为u0=v0，我们有

Uk−Ue=b1ke0 vk−Ue=b2ke0，e0=u0−Ue=v0−Ue。同样，根据命题9.2，对于每大于0，就有一些自然数），然后



此外，还有一个向量e0=e0（k），这样

ke0k=1和，

这意味着陈述（2）。

综上所述，我们发现在研究新的迭代方法时，我们必须处理以下两个问题：

1. 给出了矩阵B的迭代方法，确定该方法是否收敛。这涉及到确定是ρ（b）<1，还是等效地确定是否存在子矩阵范数，如kbk<1。根据命题8.11，这意味着i−b是可逆的（因为k−b k=k b k，命题8.11适用）。
2. 给出了两种收敛迭代方法，并进行了比较。迭代法的速度较快，其矩阵的谱半径较小。

我们现在讨论三种求解线性系统的迭代方法：

1. 雅可比方法
2. 高斯-赛德尔方法
3. 放松法。

## 9.3 Jacobi、Gauss–Seidel和Relaxy方法说明

本节描述的方法是以下方案的实例：给定一个线性系统ax=b，具有可逆性，假设我们可以将a写成

A=m-n，

其中m是可逆的，并且“易于可逆”，这意味着m接近于一个对角线或三角形矩阵（可能是分块的）。那么au=b等于

mu=nu+b，

也就是说，

U=m−1nu+m−1b。

因此，我们处于前面章节描述的情况中，b=m−1n，c=m−1b。事实上，由于a=m−n，我们有

b=m−1n=m−1（m−a）=i−m−1a，（）

这表明i−b=m−1a是可逆的。与矩阵b=m−1n相关的迭代方法由下式给出：

UK+1=m−1nuk+m−1b，k≥0，（†）

从任意向量u0开始。从实际的角度来看，我们不求m，而是迭代求解系统。

muk+1=nuk+b，k≥0。

不同的方法对应于从a中选择m和n的不同方法。前两种方法选择m和n作为a的不相交子矩阵，但松弛方法允许m和n重叠。

为了描述m和n的各种选择，用三个子量d，e，f，as来写a是很方便的。

A=D−E−F，

如果d中的唯一非零项是a中的对角线项，e中的唯一非零项是对角线下a中的项，f中的唯一非零项是对角线上a中的项。更明确地说，如果

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

 

0 0 0·····−1N−1 0··

γ

0 0 0 0···0安

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  网络错误  网络错误 | 网络错误  网络错误  网络错误  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. = D
  2. =e+f，

因此（），

b=m−1n=d−1（e+f）=i−d−1a。

作为一个符号，我们让

j=i−d−1a=d−1（e+f）

这就是雅各比矩阵。相应的方法，雅可比迭代法，使用递推计算序列（uk）。

UK+1=D−1（E+F）UK+D−1b，K≥0。

在实践中，我们迭代求解系统

duk+1=（e+f）uk+b，k≥0。

如果我们写），我们迭代地解决以下系统：

A11UK1+1=−A12UK2−A13U3K·····−A1NUKN+B1 A22UK2+1=−A21UK1−A23U3K···−A2NUKN+B2

…………

a n a−nn1n−1ku+1kn+1−1=−a a n−n111uukk1−a····n2uk2−an−1···n−2ukn−2−ann−1ukn−1−an−1nunk++bbnn−1 un

在Matlab中，雅可比迭代的一个步骤是通过以下函数实现的：

函数v=jacobi2（a，b，u）n=size（a，1）；v=zeros（n，1）；

对于i=1:n v（i，1）=u（i，1）+（-a（i，：）\*u+b（i））/a（i，i）；

结束

要运行m迭代步骤，请运行以下函数：

函数u=Jacobi（a，b，u0，m）

u=u0；对于j=1:m

u=雅各比2（a，b，u）；结束

结束

例9.1。考虑线性系统

.

我们立即检查解决方案

x1=11，x2=-3，x3=7，x4=-4.

很容易看出雅可比矩阵是

.

经过10次雅可比迭代，我们找到了近似解。

x1=10.2588，x2=-2.5244，x3=5.8008，x4=-3.7061。

经过20次迭代，我们找到了近似解

x1=10.9110，x2=-2.9429，x3=6.8560，x4=-3.9647。

经过50次迭代，我们找到了近似解

x1=10.9998，x2=-2.9999，x3=6.9998，x4=-3.9999，

经过60次迭代，我们找到了近似解

x1=11.0000，x2=-3.0000，x3=7.0000，x4=-4.0000，

最多可更正四位小数。

可以证明（见问题9.6），j的特征值是

，

所以j=b的光谱半径是

.

根据定理9.3，雅可比方法收敛于这个例子的矩阵。

观察到我们可以尝试使用第二个方程的in-solution的新值来“加速”该方法，更一般地说，在第i个方程的uki+1的求解中使用而不是。这一观察结果导致

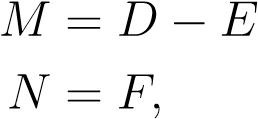
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 网络错误 |  |  |  |  |  |  |  |
| 网络错误  网络错误 | 网络错误  网络错误  网络错误 |  |  |  | 网络错误  网络错误 | 网络错误 | 网络错误 |
|  | 网络错误  网络错误 |  | 网络错误 | 网络错误 |  |  |  |

以矩阵形式写的

duk+1=euk+1+fuk+b。

因为d是可逆的，e是下三角的，所以矩阵d-e是可逆的，所以上面的方程等于

Uk+1=（d−e）−1fuk+（d−e）−1b，k≥0。上述对应于选择m和n为



矩阵b由

b=m−1n=（d−e）−1f。

由于m=d−e是可逆的，我们知道i−b=m−1a也是可逆的。

我们刚才描述的方法是高斯-赛德尔迭代法，矩阵B称为高斯-赛德尔矩阵，用l1表示，l1=（d−e）−1f。

高斯-赛德尔方法的一个优点是只需要雅可比方法使用的一半内存，因为我们只需要



计算Uki+1。我们还表明，在某些重要情况下（例如，如果a是三对角矩阵），高斯-赛德尔方法比雅可比方法收敛得更快（在这种情况下，两者同时收敛或发散）。

在matlab中，高斯-赛德尔迭代的一个步骤是通过以下函数实现的：函数u=gaussseidel3（a，b，u）n=size（a，1）；对于i=1:n u（i，1）=u（i，1）+（-a（i，：）\*u+b（i））/a（i，i）；结束

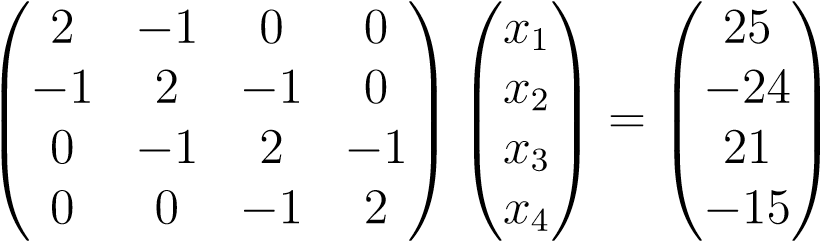
值得注意的是，与Jacobi2的唯一区别是，在赋值的两边使用相同的变量u。要运行m迭代步骤，请运行以下函数：

函数u=GausSeidel1（a，b，u0，m）u=u0；对于j=1:m

U=高斯赛德尔3（A，B，U）；结束

结束

例9.2。考虑相同的线性系统



如例9.1所示，其解决方案是

x1=11，x2=-3，x3=7，x4=-4.

经过10次高斯-赛德尔迭代，我们找到了近似解。

x1=10.9966，x2=-3.0044，x3=6.9964，x4=-4.0018。

经过20次迭代，我们找到了近似解

x1=11.0000，x2=-3.0001，x3=6.9999，x4=-4.0000。

经过25次迭代，我们找到了近似解

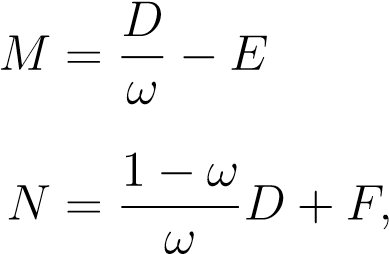
x1=11.0000，x2=-3.0000，x3=7.0000，x4=-4.0000，

最多可更正四位小数。我们观察到在这个例子中，高斯-赛德尔方法的收敛速度大约是雅可比方法的两倍。命题9.8将表明，对于三对角矩阵，高斯-赛德尔矩阵l1的光谱半径由下式给出：

ρ（l1）=（ρ（j））2，

所以我们的观察和理论是一致的。

松弛法的新成分是将矩阵d的一部分并入n中：我们用



式中，ω=06是需要适当选择的实际参数。实际上，我们在第9.4节中表明，要使松弛方法收敛，我们必须有ω∈（0,2）。注意，情况ω=1对应于高斯-赛德尔方法。

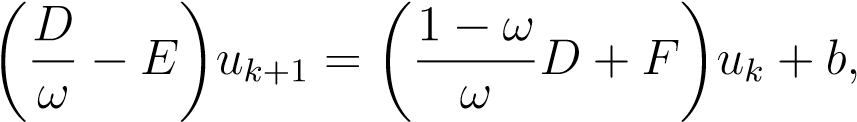
如果我们假设d的所有对角线项都不为零，那么矩阵m是可逆的。矩阵b用lω表示，称为松弛矩阵，用

.

这个数ω称为松弛参数。

当ω>1时，松弛法称为连续过松弛，简称SOR。

乍一看，松弛矩阵lω似乎比高斯-赛德尔矩阵l1复杂得多，但与松弛方法相关的迭代系统与高斯-赛德尔方法非常相似，并且非常简单。实际上，与弛豫方法相关联的系统由下式给出：



相当于

（d-ωe）uk+1=（1-ω）d+ωf）uk+ωb，

而且可以写

duk+1=duk−ω（duk−euk+1−fuk−b）。

明确地说，这就是系统

A11UK1+1=A11UK1−ω（A11UK1+A12UK2+A13U3K+····+A1N−2UKN−2+A1N−1UNK−1+A1UNK−B1）A22UK2+1=A22UK2−ω（A21UK1+1+A22UK2+A23UK3+····+A2N−2UKN−2+A2N−1UNK−1+A2UNK−B2）

…

annukn+1=annukn−ω（an11uk+1+an2uk2+1+·····+ann−2ukn+1−2+ann−1unk+1−1+annukn−bn）。

在matlab中，松弛迭代的一个步骤是通过以下函数实现的：

函数u=relax3（a，b，u，omega）n=size（a，1）；对于i=1:n u（i，1）=u（i，1）+omega\*（-a（i，：）\*u+b（i））/a（i，i）；结束

观察到函数relax3是通过在表达式（−a（i，：）u+b（i））/a（i，i）前面插入ω从函数gausseidel3获得的。要运行m迭代步骤，请运行以下函数：

函数u=松弛（a，b，u0，omega，m）u=u0；对于j=1:m

u=松弛3（a，b，u，omega）；结束

结束

例9.3。考虑与实施例9.1和9.2中相同的线性系统，其解为

x1=11，x2=-3，x3=7，x4=-4.

在ω=1.1的10次松弛迭代后，我们得到了近似解。

x1=11.0026，x2=-2.9968，x3=7.0024，x4=-3.9989。

经过10次ω=1.2的迭代，我们得到了近似解。

x1=11.0014，x2=-2.9985，x3=7.0010，x4=-3.9996。

经过10次ω=1.3的迭代，我们得到了近似解。

x1=10.9996，x2=-3.0001，x3=6.9999，x4=-4.0000.

经过10次ω=1.27的迭代，我们得到了近似解。

x1=11.0000，x2=-3.0000，x3=7.0000，x4=-4.0000，

最多可更正四位小数。我们观察到在这个例子中，ω=1.27的松弛方法比高斯-赛德尔方法收敛得更快。这一观察结果将由提案9.10予以确认。

需要做的是找到确保松弛法（和高斯-赛德尔法）收敛的条件，即：

1. 求ω的条件，即区间i r，使ω∈i表示ρ（lω）<1；证明ω∈（0,2）是一个必要条件。
2. 找出ω∈i是否存在某个最优值ω0，以便

ρ（lω0）=infρ（lω）。ωi

我们将在下一节中部分回答上述问题。

也可以通过使用A=D−E−F形式的块分解来扩展本节的方法，其中D、E和F由块组成，D是可逆的块对角矩阵。见图9.1。

D

D

D

D

E

E

E

F

F

F

1

1

1

2

2

2

3

3

3

4

图9.1：块分解的示意图a=d−e−f，其中d=4i=1di，e=3i=1ei，f=3i=1fi。

## 9.4高斯-赛德尔方法和松弛方法的收敛性

我们从与（复）厄米特正定矩阵（a=m−n）相关的迭代方法的收敛性的一般准则开始，然后将此结果应用于松弛方法。

提案9.5。设A为任意厄米特正定矩阵，写为

A=m-n，

其中m是可逆的。那么m+n是厄米提安，如果它是正定的，那么

ρ（m−1n）<1，

使迭代法收敛。

9.4。方法的收敛性

证据。因为m=a+n，a是厄米提安，a=a，所以我们得到

M+N=A+N+N=A+N+N=M+N=（M+N），

这表明M+N确实是赫米特人。因为a是厄米特正定的，函数

V 7→（V AV）1/2

从cn到r是一个向量范数k k，让k k也表示它的从属矩阵范数。我们证明了

km−1nk<1，

定理9.1证明了ρ（m−1n）<1。按定义

，

这导致我们评估kmwv−，ma−1=avak，whenakv=km=1−。如果我们写，我们有w=m−1a v，利用kvk=1，v=a−1的事实。

kv−wk2==（kvk−2 W）v a a w（v−−ww）av+w aw

-

=1−w m w mw+w aw=1−w（m+n）w。

既然我们假设m+n是正定的，如果w 6=0，那么w（m+n）w>0，我们得出结论

如果kvk=1，则kv−m−1avk<1。

最后，函数

V 7→kV−m−1avk

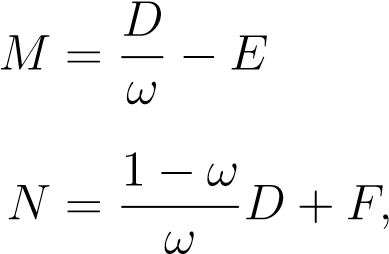
作为连续函数的一个组成部分，它是连续的，因此它在紧子集v∈cn kvk=1上达到最大值，这证明了

，

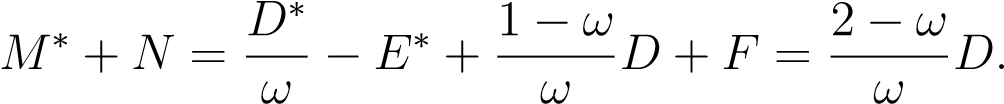
并完成证明。

现在，和前面的部分一样，我们假设a写为a=d−e−f，d可逆，可能是块形式。下一个定理提供了松弛方法收敛（因此，高斯-赛德尔方法收敛）的充分条件（这也是必要的）。这个定理被称为奥斯托夫斯基-赖希定理。

定理9.6。如果a=d−e−f是厄米特正定的，如果0<ω<2，则松弛法收敛。这也适用于证明的块分解。回想一下，对于松弛法，a=m-n



因为d=d，e=f（因为a是厄米提安的），ω=06是真的，我们有



如果d由a的对角项组成，那么从7.8节我们知道这些项都是正的，并且由于ω∈（0,2），我们看到矩阵（（2−ω）/ω）d是正定的。如果d由a的对角块组成，因为a是正的、确定的，通过选择向量z，通过为d的每一块选择一个非零向量并填充零，我们可以看到d的每一块都是正的，因此d本身是正的。因此，在所有情况下，m+n都是肯定的，我们用命题9.5得出结论。

备注：如果我们允许参数ω为非零复数ω∈c，怎么办？在这种情况下，我们

但是，

，

因此松弛法也收敛于

|ω−1<1.

如果ω为实，则此条件减小到0<ω<2。

不幸的是，定理9.6不适用于雅可比方法，但在特殊情况下，可以用命题9.5证明其收敛性。在正的一面，如果矩阵严格地是对角占优的列（行），那么可以证明雅可比方法和高斯-赛德尔方法都是收敛的。当ω∈（0,1）时，松弛法也收敛，但这不是一个非常有用的结果，因为ω>1的收敛速度通常会加快。

我们现在证明，除了a和d是可逆的之外，在没有假设a=d−e−f的情况下，为了使松弛方法收敛，我们必须有ω∈（0,2）。

9.5。三对角矩阵的收敛方法

提案9.7.给定任意矩阵a=d−e−f，且a和d可逆，对于任何ω=06，我们得到

ρ（lω）≥ω−1，

哪里。因此，松弛法（可能是分块）

除非ω∈（0,2），否则不收敛。如果我们允许ω是复数，那么我们必须

|ω−1<1

使松弛法收敛。

证据。观察Lω特征值的积λ1···········λn等于Det（Lω），由下式得出：

.

由此得出ρ（lω）≥λ1········································

如果ω∈c，证明是相同的。

## 9.5三对角矩阵的雅可比方法、高斯-赛德尔方法和松弛方法的收敛性

我们现在考虑的情况是，A是一个三对角矩阵，可能是按块划分的。在这种情况下，我们得到了关于j和lω谱半径的精确结果，从而得到了这些方法的收敛性。我们还得到了一些关于这些方法收敛速度的信息。我们从ω=1开始，这在技术上更容易处理。下面的命题给出了雅可比矩阵和高斯-赛德尔矩阵的光谱半径ρ（j）和ρ（l1）之间的精确关系。

提案9.8.设A为三对角矩阵（可能按块）。如果ρ（j）是雅可比矩阵的谱半径，而ρ（l1）是高斯-赛德尔矩阵的谱半径，那么我们得到ρ（l1）=（ρ（j））2。

因此，雅可比方法和高斯-赛德尔方法都是同时收敛或同时发散的（即使当a是由块构成的三对角线）；当它们收敛时，高斯-赛德尔方法比雅可比方法收敛得更快。

证据。我们从一个初步的结果开始。让一个（礹）具有一个三对角矩阵的分块形式

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

Det（A（μ））=Det（A（1）），μ=06。为了证明这一事实，形成块对角矩阵

p（μ）=diag（μi1，μ2i2，…，μpip）

其中，ij是与AJ块具有相同维度的单位矩阵。那么很容易看出

A（μ）=P（μ）A（1）P（μ）−1，

因此，Det（A（μ））=Det（P（μ）A（1）P（μ）−1）=Det（A（1））。

由于雅可比矩阵是j=d−1（e+f），j的特征值是特征多项式的零。

Pj（λ）=det（λi−d−1（e+f）），

因此，它们也是多项式的零

qj（λ）=det（λd−e−f）=det（d）pj（λ）。

同样，由于高斯-赛德尔矩阵是l1=（d−e）−1f，特征多项式的零点

pl1（λ）=det（λi−（d−e）−1f）

也是多项式的零

ql1（λ）=det（λd−λe−f）=det（d−e）pl1（λ）。

由于a=d−e−f是三对角（或按块划分的三对角），因此λ2d−λ2e−f也是三对角（或按块划分的三对角），并且通过使用我们对μ=λ=06的初步结果，我们得到

ql1（λ2）=det（λ2d−λ2e−f）=det（λ2d−λe−λf）=λnqj（λ）。

通过连续性，上述方程也适用于λ=0。但我们推断：

1. 对于任何β=06，如果β是l1的特征值，那么β1/2和−β1/2都是j的特征值，其中β1/2是β的复平方根之一。

9.5。三对角矩阵的收敛方法

1. 对于任何α=06，如果α和−α都是j的特征值，那么α2是l1的特征值。

上面立即暗示了ρ（l1）=（ρ（j））2。

现在我们考虑更一般的情况，其中ω是（0,2）中的任何实数。

提案9.9.假设A是一个三对角矩阵（可能是分块的），并假设雅可比矩阵的特征值都是实的。如果ω∈（0,2），那么雅可比方法和松弛方法都会同时收敛或同时发散（即使当a为三对角块时）。当它们收敛时，函数ω7→ρ（lω）（对于ω∈（0,2））的唯一最小值等于ω0−1

，

其中1<ω0<2，如果ρ（j）>0。我们还有ρ（l1）=（ρ（j））2，如前所述。

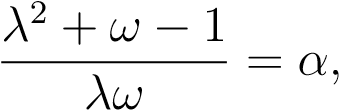
证据。这是一个非常技术性的证明，可以在serre[151]和ciarlet[41]中找到。在前面的命题的证明中，我们首先证明矩阵lω的特征值是多项式的零。

，

式中，plω（λ）是lω的特征多项式。然后利用9.8号提案的初步事实，很容易证明

，

对于所有的λ∈c，λ=06。这一次我们不能把上述方程推广到λ=0。这导致我们考虑这个方程

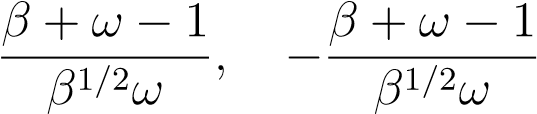


相当于

λ2−αωλ+ω−1=0，

对于所有λ=06。自λ=06以来，上述等价性不适用于ω=1，但这不是问题，因为在前面的命题中已经考虑了ω=1的情况。然后我们可以展示以下内容：

1. 对于任何β=06，如果β是lω的特征值，那么



是j的特征值。

1. 对于每个α=06，如果α和−α是j的特征值，那么μ+（α，ω）和μ−（α，ω）是lω的特征值，其中μ+（α，ω）和μ−（α，ω）是方程λ2−αωλ+ω−1=0根的平方。

接下来是

ρ（lω）=maxj max（+（α，ω），−（α，ω），

λp（λ）=0

既然我们假设j有实根，我们就可以研究它的函数。

m（α，ω）=max（α，ω），（α，ω），

其中α∈R和ω∈（0,2）。实际上，因为m（−α，ω）=m（α，ω），只需要考虑α≥0的情况。

注意，对于α=06，方程式的根

是

.

反过来，这会导致考虑方程式的根。

，

哪些是

α=06。既然我们有

和

，

这些根是

.

−−−

注意ω0（α）的表达式正是我们命题中的表达式！其余的证明包括通过考虑α的各种情况来分析函数m（α，ω）的变化。最后，我们发现ω0（ρ（j））得到了ρ（lω）的最小值。细节很冗长，我们省略了。读者将在Serre[151]和Ciarlet[41]中找到完整的证据。

9.6。总结

结合定理9.6和9.9的结果，我们得到了下面的结果，给出了矩阵j、l1和lω的光谱半径的精确信息。

提案9.10。设A是一个三对角矩阵（可能是分块的），它是厄米特正定的。然后，雅可比方法、高斯-赛德尔方法和松弛方法都收敛于ω∈（0,2）。有一个独特的最佳松弛参数

，

这样的话

ρ（lω0）=infρ（lω）=ω0−1。

0＜ω＜2

此外，如果ρ（j）>0，则

ρ（lω0）<ρ（l1）=（ρ（j））2<ρ（j），

如果ρ（j）=0，那么ω0=1，ρ（l1）=ρ（j）=0。

证据。为了应用9.9号命题，我们必须检查j=d−1（e+f）是否具有实特征值。但是，如果α是j的任何特征值，如果u是任何对应的特征向量，那么

d−1（e+f）u=αu

意味着

（e+f）u=αdu，

由于a=d−e−f，上述结果表明（d−a）u=αdu，即，

au=（1−α）du.

因此，u au=（1−α）u du，

由于a和d是厄米特正定的，当u=06时，我们得到u au>0和u du>0，这证明了αr，其余的都符合定理9.6和9.9。

注：当最佳松弛参数未知时，最好高估而不是低估松弛参数。

## 9.6总结

本章的主要概念和结果如下：

* 迭代法。将a拆分为a=m−n。
* 向量或矩阵序列的收敛性。
* 矩阵b的幂次收敛到零的一个准则（以谱半径ρ（b）.bk为单位）
* 光谱半径ρ（b）作为序列极限（Kbkkk1/k）的特征。
* 迭代法收敛的一个准则。
* 迭代方法的渐近行为。
* 分裂（和索拉）。作为a=d−e−f，以及雅可比、高斯-赛德尔和松弛的方法
* 雅可比矩阵，j=d−1（e+f）。
* 高斯-赛德尔矩阵，l1=（d−e）−1f。
* 松弛矩阵，lω=（d−ωe）−1（（1−ω）d+ωf）。
* 正定迭代法的收敛性：当a=m−n为Hermitian时的一般结果
* 雅可比方法、高斯-赛德尔方法和松弛方法收敛的一个充分条件。奥斯托夫斯基-赖希定理：a是厄米特正定，ω∈（0,2）。
* 雅可比方法、高斯-赛德尔方法和松弛方法收敛的一个必要条件：ω∈（0,2）。
* 三对角矩阵的情况（可能是按块）。Jacobi方法和Gauss-Seidel方法同时收敛或发散，并比较了ρ（j）和ρ（l1）：ρ（l1）=（ρ（j））2的光谱半径。
* 三对角厄米正定矩阵的情况（可能是分块的）。雅各比、高斯-塞德尔和放松的主题方法都在汇聚。
* 在上述情况下，有一个唯一的最佳松弛参数，其中（l1）=（ρ（j））2<ρ（j）（如果ρ（j）=0）.6ρ（lω0）<

## 9.7问题

问题9.1。考虑矩阵

.

9.7。问题

证明ρ（j）=0和ρ（l1）=2，所以

ρ（j）<1<ρ（l1），

其中j是雅各比矩阵，l1是高斯-赛德尔矩阵。

问题9.2。考虑矩阵

.

证明ρ（j）=√5/2和ρ（l1）=1/2，所以

ρ（l1）<ρ（j），

其中j是雅各比矩阵，l1是高斯-赛德尔矩阵。

问题9.3。考虑以下线性系统：

.

1. 用高斯消去法求解上述系统。
2. 利用Jacobi、Gauss–Seidel的方法计算向量10的序列，并松弛ω：ω=1.1,1.2，…，1.9的以下值。在所有情况下，初始向量都是u0=（0,0,0,0）。

问题9.4。回想一下，复数或实数n×n矩阵a严格地是行对角占优的if。

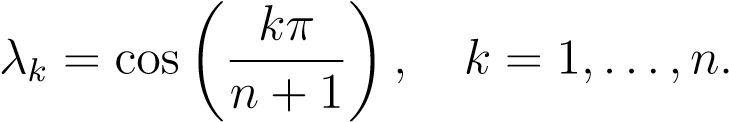
1. 证明了如果a是严格的行对角占优，则雅可比方法收敛。
2. 证明了如果A是严格的行对角占优，则高斯-赛德尔方法收敛。

问题9.5。证明9.5号命题的逆命题成立。也就是说，如果a是一个厄米正定矩阵，写为a=m−n，m可逆，如果厄米坦矩阵m+n是正定的，如果ρ（m−1n）<1，则a是正定的。

问题9.6.考虑以下三对角n×n矩阵：

.

（1）证明雅可比矩阵j的特征值由



暗示。首先证明雅可比矩阵是

.

那么j的特征值和特征向量就是方程组的解。

Y0＝0

yk+1+yk−1=2λyk，k=1，…，n yn+1=0。

众所周知，上述复发的一般解决方案如下：

，

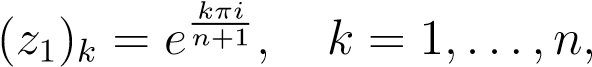
（α，β=0）6，其中z1和z2是方程的零。

z2−2λz+1=0.

因此，z2=z1−1和z1+z2=2λ。边界条件y0=0生成α+β=0，所以），边界条件yn+1=0生成

.

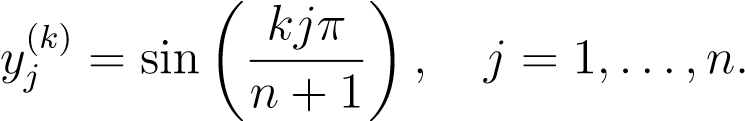
我们可以假设z1的n个可能值（z1）k由下式给出：



找到

2λk=（z1）k+（z1）−k 1.

证明与特征值λk相关联的特征向量（）由下式给出：



（2）求谱半径ρ（j）、ρ（l1）和ρ（lω0），作为h=1/（n+1）的函数。

第十章

# 二元空间，二元性

## 10.1双空间E和线性形式

在第3.8节中，我们定义了向量空间e的对偶空间e=hom（e，k），并证明了有限维向量空间的对偶基的存在性。

在本章中，我们更深入地研究了空间E与其双空间E之间的联系。正如我们即将看到的，每一个线性映射f:e→f都会产生一个线性映射f>：f→e，结果表明，在适当的基础上，f>的矩阵是f的矩阵的转置。因此，对偶空间的概念提供了与tra相关的现象的概念解释。n位置。

但是它做得更多，因为它允许我们把一个线性方程看作双空间E的一个元素，从而把E的子空间看作一组线性方程的解，反之亦然。子空间与线性形式集之间的关系是对偶性的本质，一个经常被松散使用的术语，但可以精确地作为给定向量空间e的子空间集与其对偶e的子空间集之间的双射。在这种对应关系中，e的子空间v产生e的子空间v 0，由在v上消失的所有线性形式组成（即，v中所有输入的值为零）。考虑下一组r3中的两个“线性方程”，

1. −Y+Z=0 x−Y−Z=0，

让我们找出它们的公共解（x，y，z）的集合v∈r3。通过从第一个方程中减去第二个方程，我们得到2z=0，并且通过添加这两个方程，我们发现2（x-y）=0，因此解的集合v由下式给出：

1. =x z=0。

这是r3的一维子空间。几何上，这是平面z=0中方程y=x的直线。

三百一十九

现在，为什么我们说上面的方程是线性的？这是因为，作为（x，y，z）的函数，两个映射f1:（x，y，z）→7 x−y+z和f2:（x，y，z）7→x−y−z都是线性的。从r3到r的所有这些线性函数的集合是一个向量空间；我们利用这个事实来形成“方程式”f1和f2的线性组合。观察子空间V的尺寸为1。环境空间的尺寸为n=3，有两个“独立”方程f1、f2，因此由m独立方程定义的子空间v的尺寸dim（v）似乎是

尺寸（V）=N−m，

这确实是一个普遍的事实。

一般来说，在RN中，线性方程是由一个n-元组（a1，…，an）∈RN决定的，该线性方程的解是由n-元组（x1，…，xn）∈RN给出的，这样a1x1+·····+anxn=0；

这些解构成了线性映射（x1，…，xn）7→a1x1+·····+anxn的核心。上述考虑假设我们在RN的规范基（e1，…，en）中工作，但是我们可以通过将“线性方程”视为从E到K场的线性映射的向量空间hom（e，k）的元素，来定义独立于基和任何维的“线性方程”。

定义10.1.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。

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10.2.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。

第3.8节解释了术语坐标表的原因。

我们在定理3.18中证明，如果（u1，…，un）是e的基，那么（）是e的基，称为对偶基。

10.1。双空间E和线性形式

如果（u1，…，un）是rn的基（更一般地说是kn），则可以显式地找到对偶基（），其中每个u i由行向量表示。例如，考虑b'ezier矩阵的列

.

由于条件定义的公式=0，因此它由行向量（λ1λ2λ3λ4）表示，以便

.

这意味着这是B4的倒数第1行。自从

，

线性形式（）对应于B4-1的行。特别地，用（1 1 1 1 1）表示。

上述方法适用于任意n。给定rn的任何基（u1，…，un），如果p是n×n矩阵，其jth列为uj，则对偶形式由矩阵p-1的第i行给出。

当e是有限维n且（u1，…，un）是e的基础时，我们注意到该族

）是双空间e（称为（u1，…，un）的双基）的基础。让我们看看双基上的线性形式\_的坐标是如何随基的变化而变化的。

设（u1，…，un）和（v1，…，vn）为e的两个基，设p=（aij）为基矩阵从（u1，…，un）变为（v1，…，vn），使

，

设p−1=（bij）为p的倒数，因此

.

由于u i（uj）=δij和vi（vj）=δij，我们得到

，

因此

，

和

.

这意味着基础从双重基础（）转变为双重基础。自从

，

我们得到

，

因此，使用矩阵p>将新坐标j表示为旧坐标\_i。如果我们使用行向量（\_，…，n）和（），我们有



与基础变化比较

，

我们注意到，这一次，线性形式\_的坐标（\_i）与基础变化方向相同。因此，我们说线性形式的坐标是协变的。由于语言的滥用，人们常说线性形式是协变的，这就解释了为什么covector这个术语也用于线性形式。

如果（e1，…，e n）是向量空间e的基础，那么，作为从e到k的线性映射，每一个线性形式f∈e都由1×n矩阵表示，也就是说，由一个行向量表示。

（λ1，…，λn）

关于e的基（e1，…，e n）和k的基（1），其中f（ei）=λi。向量u=由n×1矩阵表示，即由列向量表示。

，

10.1。双空间E和线性形式

f对u的作用，即f（u），用矩阵积表示。

.

另一方面，对于对偶基（，线性形式f是

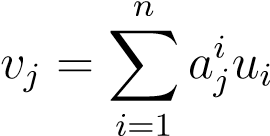
由列向量表示

.

备注：在许多使用张量的文本中，向量通常用较低的索引索引。如果是这样，则更方便的是使用上面的索引将向量X的坐标写在基（U1，…，Un）上，这样，

，

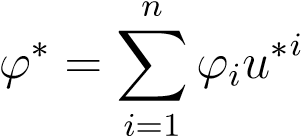
在基础的变化中，我们



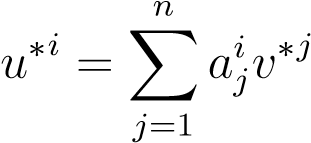
和

.

同时，线性形式被索引为高索引。然后，在双基（u 1，…，u n）上使用较低的索引将covector的坐标\_写为（\_i），以便



在基础的变化中，我们



和

.

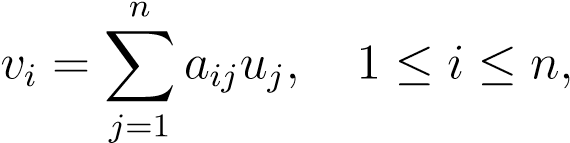
有了这些约定，求和的指标在上下两个位置分别出现一次，求和符号可以被安全地省略，这是爱因斯坦的一个技巧。例如，我们可以写



作为的缩写

.

对于爱因斯坦符号的另一个例子，如果向量（v1，…，vn）是向量（u1，…，un）的线性组合，则



那么上述方程是：

vi=Ajiuj，1≤i≤n。

因此，在爱因斯坦的表示法中，n×n矩阵（aij）用（aji）表示，这是一个（1，1）张量。

注意，有些作者认为矩阵是坐标之间的映射，在这种情况下，矩阵（aij）用（aij）表示。

## 10.2 E与E\_的配对与对偶性

给定线性形式u∈e和向量v∈e，应用u to v的结果u（v）也用hu，v i表示。这定义了一个满足以下性质的二元运算h−，−i:e×e→k：

hu 1+u 2，vi=hu 1，vi+hu 2，vi

胡，v1+v2i=hu，v1i+hu，v2i hλu，vi=λhu，vi hu，λvi=λhu，vi.

上面的恒等式意味着h−、−i是双线性映射，因为它在每个参数中都是线性的。通常称为e和e之间的正则对，鉴于上述恒等式，给定任何固定向量v e，映射evalv:e→k（v处的评价）定义如下：

evalv（u）=hu，vi=u（v）对于每个u∈e

是从e到k的线性映射，也就是说，evalv是e中的线性形式。同样，根据上述身份，地图evale:e→e，定义如下：

evale（v）=evalv，对于每个v∈e，

是一个线性映射。注意

evale（v）（u）=hu，vi=u（v），对于所有v∈e和所有u∈e。

10.2。E与E的配对与对偶性

我们将看到映射evale是内射的，当e有有限维时它是同构的。

我们现在将线性方程组v 0的概念形式化，在给定子空间v e中的所有向量上消失，并将线性方程组u e的公共解集u0的概念形式化。对偶定理（定理10.1）表明，v和v 0的维数，以及u和u0的维数，都以一种关键的方式联系在一起。在有限维上，图V 7→V 0和U 7→U0是从E的子空间到E的子空间的逆双射。

定义10.3.在给定向量空间e及其对偶e的情况下，如果hu，vi=0，则向量v e和线性形式u e是正交的。给定e的子空间v和e的子空间u，我们认为对于每个u∈u和每个v∈v，v和u是正交的，如果hu，vi=0。给出e（resp.e的一个子集u），v的正交v 0是e的子空间v 0，定义如下：

v 0=u∈e hu，vi=0，对于每个v∈v

（响应u的正交u0是e的子空间u0，定义如下：

u0=v∈e hu，vi=0，对于每个u∈u）。

子空间v 0 e也被称为v的湮灭子。被u e\_歼灭的子空间u0 e没有特殊名称。把它称为线性子空间（或线性变化）似乎是合理的。

非正式地说，v 0是在v上消失的线性方程组，u0是在u中所有线性方程的公共零点组。我们也可以用

V 0=U∈E V Keru

和U0

.

观察e0=0=（0）和0 0=e。此外，如果v1 v2 e，则

，如果U1 U2 E，那么。

实际上，如果v1 v2 e，那么对于任何一个，对于所有v∈v2，我们都有f（v）=0，因此对于所有v∈v2，f（v）=0。同样地，如果u1 u2 e，那么对于任何一个，对于所有的f（v）=0，对于所有的f∈u2，那么f（v）=0，对于所有的f∈u1，这意味着。

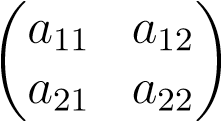
下面是一些例子。设e=m2（r），实2×2矩阵的空间，设v为矩阵所跨越的m2（r）的子空间。

.

我们会立即检查子空间v是否由该形式的所有矩阵组成。

，

也就是说，所有的对称矩阵。矩阵



在v中满足方程

A12−A21=0，

所有这些方程的标量倍数，所以v 0是e的子空间，用u（a11，a12，a21，a22）给出的线性形式表示。根据对偶定理（定理10.1），我们有

尺寸（V 0）=Dim（E）−Dim（V）=4−3=1。

上面的例子推广到e=mn（r）对于任何n≥1，但这次，考虑线性形式的空间u，断言矩阵a是对称的；这些是由n（n−1）/2方程所跨越的线性形式。

aij−aji=0，1≤i<j≤n；

注意，对角项和半个方程没有限制。

aij−aji=0，1≤i=6 j≤n

是冗余的。很容易验证i<j的方程（线性形式）是线性独立的。更精确地说，假设u是E中线性形式的空间，由线性形式u i j（a11，…，a1n，a21，…，a2n，…，an1，…，ann）=aij−aji，1≤i<j≤n所跨越。

然后，这些方程的公共解的集合u0是对称矩阵的空间s（n）。根据对偶定理（定理10.1），这个空间有维

.

我们把它作为一个练习来寻找s（n）的基础。

如果e=mn（r），考虑e中由线性形式所跨越的线性形式的子空间u。

u i j（a11，…，a1n，a21，…，a2n，…，an1，…，ann）=aij+aji，1≤i<j≤n u ii（a11，…，a1n，a21，…，a2n，…，an1，…，ann）=aii，1≤i≤n。

10.2。E与E的配对与对偶性

很容易看出这些线性形式是线性独立的，所以dim（u）=n（n+1）/2。满足上述方程的矩阵a∈mn（r）的空间u0显然是斜对称矩阵的空间斜交（n）。根据对偶定理（定理10.1），u0的维数为

.

我们把它作为一个练习来寻找歪斜（n）的基础。

对于另一个例子，e=mn（r），对于任何a∈mn（r），考虑由tr（a）=a11+a22+·········+ann给出的e中的线性形式，

称为a的迹线。e的子空间u0由所有矩阵a组成，这样tr（a）=0是尺寸n2-1的空间。我们把它作为一个练习来寻找这个空间的基础。尺寸方程

尺寸（V）+尺寸（V 0）=Dim（E）Dim（U）+Dim（U0）=Dim（E）

总是正确的（如果e是有限维的话）。这是对偶定理（定理）的一部分

10.1）。

在前面的例子中，给定矩阵a∈mn（r），断言a>a=i的方程不是线性约束。例如，对于n=2，我们有

A211+A221=1 A221+A222=1 A11A12+A21A22=0。

评论：

1. 符号v 0（分别为u0）表示e（resp）的子空间v的正交。e）的子空间u不是通用的。其他作者使用符号v（resp.u）。然而，符号v也用于表示子空间v相对于空间e上内积的正交补码，在这种情况下，v是e的子空间，而不是e的子空间（见第11章）。为了避免混淆，我们更喜欢使用表示法v 0。
2. 由于线性形式可以看作线性方程（至少在有限维中），给定e的一个子空间（甚至一个子空间）u，我们可以用

z（u）=v∈e u（v）=0，对于所有u∈u。

当然z（u）=u0，但概念z（u）可以推广到更一般的方程，即多项式方程。在这个更一般的设置中，u是一组n个变量的多项式，系数为k（其中n=dim（e））。形式z（u）的集合称为代数变种。线性形式对应于一阶齐次多项式被考虑的特殊情况。

如果v是e的子集，那么很自然地将k[x1，…，xn]中消失在v上的多项式集与v联系起来。这个集合，通常表示为i（v），有一些特殊的属性使它成为理想。如果v是e的一个线性子空间，我们自然会把注意力限制在v上消失的线性形式的空间v 0上，在这种情况下，我们识别i（v）和v 0（尽管从技术上讲，i（v）不再是理想的）。

对于任意一组多项式u k[x1，…，xn]（resp v e），i（z（u））和u（resp）之间的关系。z（i（v））和v）通常不简单，即使我们总是

u i（z（u））（分别为v z（i（v）））。

然而，当场k是代数闭的，那么i（z（u））等于理想u的根式，这是希尔伯特的著名结果，被称为nullstellensatz（见lang[106]或dummit和foote[55]）。代数变种的研究是代数几何学的主要课题，是一门美丽而又令人生畏的学科。要了解代数几何，请参见lang[106]或dummit和foote[55]。

对偶定理（定理10.1）表明，如果我们把注意力局限于线性子空间，情况会简单得多；在这种情况下

u=i（z（u））和v=z（i（v））。

我们声称e的每个子空间v v 00，e的每个子空间u的u u00。

实际上，对于任何v∈v，为了证明v∈v 00，我们需要证明u（v）=0对于所有u∈v 0。然而，v 0由所有线性形式u组成，因此u（y）=0表示所有y∈v；特别是，因为v v，u（v）=0表示所有u∈v 0，根据需要。

同样，对于任何u∈u，为了证明u∈u00，我们需要证明u（v）=0对于所有v∈u0。但是，u0由所有向量v组成，因此f（v）=0表示所有f∈u；特别是，由于u∈u，u（v）=0表示所有v∈u0，根据需要。

我们很快就会看到，在有限维中，我们有v=v 00和u=u00。

然而，即使v=v 00总是真的，当e是无限维时，u=u00也不总是真的。

在给定向量空间e和e的子空间u的情况下，根据定理3.5，u的每一个基（u i）i∈i都可以推广到e的基（u j）j∈i j，其中i j为∅。

## 10.3对偶定理

我们有以下重要的定理改编自E.Artin[6]（第1章）。

定理10.1。（对偶定理）让e是一个向量空间。以下属性保留：

1. 对于e的每一个基（ui）i∈i，坐标形式族是线性独立的。
2. 对于e的每个子空间v，我们有v 00=v。
3. 对于e的有限余维m的每个子空间v，对于e的每个子空间w，使e=v\_w（其中w是有限维m），对于e的每个基（ui）i∈i，使（u1，…，um）是w的基，该族是正交的基。

e中v的0，因此dim（v 0）=codim（v）。

此外，我们有v 00=v。

1. 对于e的有限维m的每个子空间u，e中u的正交u0是有限余维m，因此

codim（u0）=dim（u）。

此外，u00=u。

证据。（a）假设

xλiu i=0，

我爱我

对于一个家族（λi）i∈i（k中的标量）。因为（λi）i∈i有有限的支持，所以i有一个有限的子集j，因此，对于所有i∈i−j，λi=0，我们有

xλju j=0.

J·J·J

将线性形式pj∈jλj u j应用于每一个uj（j∈j），根据定义10.2，因为u i（uj）=1，如果i=j，否则为0，我们得到所有j∈j的λj=0，即所有i∈i的λi=0（根据j的定义作为支持）。因此，（是线性无关的。

V

（因此，ui）（i∈b）显然，我们有ij0是∈/i的基础。从0开始（其中v∈00v，i00u。如果j0jis与=v∅=6）中的每种线性形式正交。从00开始，那么让（v=6 Vu00i，）i∈uij 0 j∈作为0的一个00的基础。现在，我们有vj000，这样∈j（和

UJ V

）=0表示所有i∈i，因此。然而，）=1，这与uj0与v 0中的所有线性形式都是正交的这一事实相矛盾。因此，v=v 00。

1. 设j=i−1，…，m。每一个线性形式f∈v 0与每一个uj正交，对于j j，因此，f（uj）=0，对于所有j j。对于这样一个线性形式f∈v 0，让

.

我们有g（ui）=f（ui），对于每个i，1≤i≤m。此外，根据定义，g在所有uj上消失，其中j∈j。因此，f和g在（ui）i∈i of e，and so，g=f。

这表明（）产生了v 0，由于它也是一个线性独立的族，

）是V 0的基础。很明显，dim（v 0）=codim（v），根据（b）部分，我们得到v 00=v。

1. 让（）作为u的基础。注意地图h:e→km的定义如下：



对于每一个v∈e，都是一个线性映射，其核kerh精确地为u0。

提案5.11，e≈ker（h）imh=u0 imh，

由于dim（imh）≤m，我们推断u0至多是有限余维e的一个子空间，通过（c），我们得到dim（u00）=codim（u0）≤m=dim（u）。然而，很明显u u00，这意味着dim（u）≤dim（u00），所以dim（u00）=dim（u）=m，我们必须有u=u00。

定理10.1的（a）部分表明

尺寸（e）≤尺寸（e）。

当e是有限维n且（u1，…，un）是e的基础时，通过（c）部分，族

）是双空间e的基础，称为（u1，…，un）的双基础。定理3.18也直接证明了这一事实。

定义从e的子空间到e的子空间的函数e（e表示方程）和从e的子空间到e的子空间的函数z（z表示零），方法是

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

（z\_e）（v）=v 00=v

（e\_z）（u）=u00=u，

其中v是e的有限余维的子空间，u是e的有限维的子空间，因此映射e和z是这些子空间之间的反向双射。这些映射在e的有限余维子空间和e的有限维子空间之间建立了对偶性。特别是，如果e是有限维，那么维m的每个子空间v e是线性形式空间（方程）v 0的公共零点集，它的维数为n−m。这证实了我们对由一组线性方程定义的子空间的维数所作的声明。

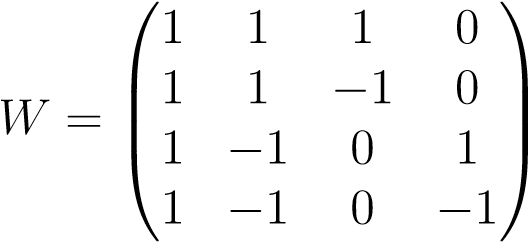
注意这个双射不延伸到无限维的e的子空间。

当e是无限维时，对于e的每一个基（u i）i∈i，坐标形式的族（u i）i∈i绝不是e的基。它是线性独立的，但它“太小”不能产生e。例如，如果e=r（n），其中n=0,1,2，…，则表示e中向量的非零坐标和的映射f:e→r是线性形式，但很容易看出它不能表示为坐标形式的线性组合。因此，当e是无限维时，e和e不是同构的。

假设v是维m的rn的子空间，而（v1，…，vm）是v的基础。为了找到v 0的基，我们首先将（v1，…，vm）扩展到rn的基（v1，…，vn），然后通过定理10.1的（c）部分，我们知道（）是v0的基。例如，假设v是两个线性无关向量所跨越的r4的子空间。

，

r4中haar基的前两个向量。haar矩阵的四列



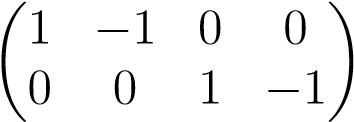
构成r4的基础，w的倒数由

.

因为双基（）由w-1的行给出，w-1的最后两行，

，

构成V 0的基础。我们还通过因子1/2重新缩放获得了一个基，因此行向量给出的线性形式



形成V 0的基础，即线性形式（线性方程）在子空间V上消失的空间。

我们描述的求v 0的方法需要先扩展v的基，然后求逆矩阵，但还有一种更直接的方法。实际上，假设a是n×m矩阵，其列是v的基向量（v1，…，vm）。那么，由行向量表示的线性形式u属于v 0 iff u v i=0，i=1，…，m iff。

UA＝0

敌我识别

A>U>=0。

因此，我们需要做的就是找到>的空空间的基础。这可以通过将矩阵约简为约简行梯队形式（RREF）来非常有效地实现；参见第节

7.10。

现在让我们考虑一个问题，找到由方程定义的Rn中超平面h的基。

c1x1+·····+cnxn=0.

更准确地说，如果u（x1，…，xn）是u（x1，…，xn）=c1x1+······+cnxn给出的（rn）中的线性形式，那么超平面h是u的核心。当然，我们假设一些Cj是非零的，在这种情况下，线性形式u跨越（rn）的一维子空间u，并且u0=h具有尺寸n-1。

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

由于u不是一个相同为零的线性形式，所以有一个最小的正指数j≤n，使得cj=06，所以我们的线性形式实际上是u（x1，…，xn）=cjxj+······+cnxn。我们

j1 0 0…1 0 0…0

J 0 0…0 C/C C/C…C/C\_

观察删除j行得到的（n-1）×n-1）矩阵是单位矩阵，因此上述矩阵的列是线性无关的。一个简单的计算也表明，线性形式u（x1，…，xn）=cjxj+·····+cnxn在上述矩阵的每一列上都消失了。对于R6中的一个具体例子，如果u（x1，…，x6）=x3+2x4+3x5+4x6，我们得到方程的超平面h的基础。

x3+2x4+3x5+4x6=0

由以下矩阵给出：

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

相反，给定一个超平面h，在rn中，作为n-1线性向量的跨度。

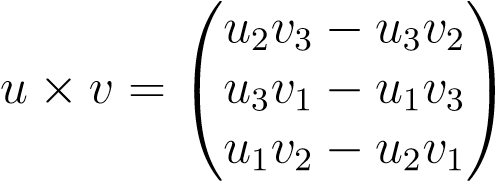
（u1，…

和

是两个线性无关的向量，那么

，

u和v的叉积u×v由



是一个解决方案。

下面是另一个例子，说明定理10.1的幂。设e=mn（r），并考虑断言矩阵a∈mn（r）每一行中的项之和等于同一个数的方程。我们有n-1方程

，

很容易看出它们是线性独立的。因此，由上述线性形式（方程）所跨越的e中线性形式的空间u的维数为n−1，而将所有这些方程化的矩阵的空间u0的维数为n2−n+1。找到这个空间的基础并不那么明显。

现在我们将确定向量空间e与其双e之间的关系。

提案10.2.设e为向量空间。以下属性保留：

1. 线性映射evale:e→e定义如下：

evale（v）=evalv，对于所有v∈e，

也就是说，evale（v）（u）=hu，vi=u（v）对于每个u∈e，都是内射的。

1. 当e为有限维n时，线性映射evale:e→e是同构（称为规范同构）。

证据。（a）设（ui）i∈i为e的基，设v=pi∈i viui。如果evale（v）=0，那么特别是eval）=0表示全部，并且自

埃瓦

对于所有i∈i，我们有vi=0，也就是说，v=0，表明evale:e→e是内射的。

如果e是有限维n，根据定理10.1，对于每个基（u1，…，un），族

）是双空间e的基础，因此族（）是双空间e的基础。这表明dim（e）=dim（e）=n，由于（a）部分，我们知道evale:e→e是内射的，实际上evale:e→e是双射的（因为一个内射映射将一个线性独立的族传递给一个线性独立的族，在一个维数为n的向量空间中，一个li几乎独立的n向量族是一个基础，见命题

3.6）。

当一个向量空间E有无限维时，E及其双e永远不同构。

当e是有限维且（u1，…，un）是e的基础时，考虑到正则同构evale:e→e，双元的基础（）用

（U1，…，联合国）。

命题10.2可以用配对的方式非常有效地重新表述，这是庞特贾金在1931年发现的一个非常有用的概念（改编自E.Artin[6]，第1章）。在一个k域上给定两个向量空间e和f，我们认为，如果对于每一个v∈v，映射u 7（u，v）（从e到k）是线性的，对于每一个u∈e，映射v 7→（u，v）（从f到k）是线性的，那么函数\_：e×f→k是双线性的。

定义10.4.在两个向量空间e和f超过k的情况下，e和f之间的配对是双线性映射，即：e×f→k。这种配对是非退化的iff。

1. 对于每一个u∈e，如果所有v∈f的（u，v）=0，则u=0，并且
2. 对于每一个v∈f，如果所有u∈e的（u，v）=0，则v=0。

一个配对：e×f→k通常用h−、−i:e×f→k表示。例如，前面定义的映射h−、−i:e×e→k是一个非退化配对（使用命题10.2中（a）的证明）。如果e=f和k=r，e上的任何内积都是非退化配对（因为内积是肯定的）；见第11章。

10.4。超平面和线性形式

给出了一个配对，我们可以定义两个映射，l\_：e→f和r\_：f→e如下：对于每个u∈e，我们在f\_中定义线性形式l\_（u），这样

l\_（u）（y）=\_（u，y），对于每个y∈f，

对于每一个v∈f，我们定义e中的线性形式r\_（v），这样

每x∈e，r（v）（x）=（x，v）。

我们有以下有用的建议。

提案10.3.给定两个向量空间e和f over k，对于e和f之间的每一个非退化配对，映射l\_：e→f和r\_：f→e\_是线性的和内射的。此外，如果e和f有有限的尺寸，那么这个尺寸是相同的，l\_：e→f和r\_：f→e是双射。

证据。因为配对是双线性的，所以图l\_：e→f和r\_：f→e\_是线性的。如果l\_（u）=0（零形式），则

对于每一个v∈f，

既然\_是非退化的，u=0。因此，l\_：e→f是注射的。同样，R\_：F→E是注射剂。当f有有限维数n时，我们已经看到f和f有相同的维数。因为l\_：e→f是内射的，我们有m=dim（e）≤dim（f）=n。同样的论点也适用于e，因此n=dim（f）≤dim（e）=m。但是，dim（e）=dim（f），l\_：e→f和r\_：f→e是双射。

当e有有限维时，非简并配对h−，−i:e×e→k给出了e与e之间存在自然同构的另一个证明。当e=f时，由e上的内积引起的非退化配对产生e和e之间的自然同构（见第11.2节）。

有趣的非退化配对出现在外代数中。我们现在展示超平面和线性形式之间的关系。

## 10.4超平面和线性形式

实际上，下面的命题10.4是从定理10.1的（c）和（d）部分得出的，但我们认为给出一个更直接的证明也是有趣的。

提案10.4.设e为向量空间。以下属性保留：

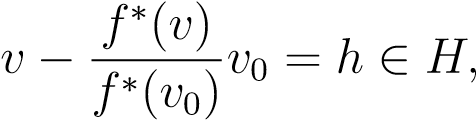
1. 给定任意非零线性形式f∈e，其核h=kerf是超平面。
2. 对于e中的任何超平面h，都有一个（非空）线性形式f∈e，这样h=kerf。
3. 给定e中的任意超平面h和任意（非空）线性形式f∈e，使得h=kerf，对于每个线性形式g∈e，h=kerg iff g=λf，对于k中的某些λ=06。

证据。（a）如果f∈e为非空，则存在一些向量v0∈e，使得f（v0）=06。让

H=切口。对于每一个v∈e，我们有

.

因此，

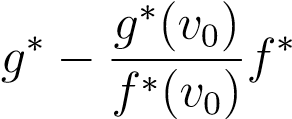


和

，

也就是说，e=h+kv0。另外，由于f（v0）=06，我们得到v0∈/h，即h kv0=0。因此，e=h kv0，h是超平面。

1. 如果h是一个超平面，对于某些v0∈/h，e=h kv0，那么每个v∈e都可以用一种独特的方式写成v=h+λv0。因此，有一个定义明确的函数f：e→k，因此，对于每个v=h+λv0，f（v）=λ。作为一个简单的练习，我们将验证f是一个线性形式。由于f（v0）=1，线性形式f为非空。而且，根据定义，很明显，λ=0 iff v∈h，即切口=h。
2. 设h为e中的超平面，设f∈e为任意（非空）线性形式，使得h=kerf。显然，如果g=λf对于某些λ=06，则h=kerg。相反，假设对于某些非空线性形式g，h=kerg。从（a），我们得到e=h kv0，对于一些v0，这样f（v0）=06和g（v0）=06。然后，观察



是一个在h上消失的线性形式，因为f和g都在h上消失，但也在kv0上消失。因此，g=λf，与

.

作为练习，我们将向量空间e的每个子空间v=6e，都是包含v的所有超平面的交集。我们现在考虑线性映射和矩阵转置的概念。

## 10.5线性映射和矩阵的转置

给定一个线性映射f:e→f，可以定义一个具有一些有趣特性的映射f>：f→e。

定义10.5.给定线性映射f:e→f，f的转置f>：f→e是定义如下的线性映射：

f>（v）=v f，对于每个v∈f，

如下图所示：

f

fe>（bb bbbb/f v\_

五）

K

等效地，线性映射f>：f→e的定义如下：

hv，f（u）i=hf>（v），ui，

对于所有u e和所有v f。

很容易验证以下属性是否有效：

1. +g）>=f>+g>
2. \_f）>=f>g>id.

注意（g f）>=f>g>右侧的成分倒转。

方程式（g f）>=f>g>包含以下有用的命题。

提案10.5.如果f:e→f是任何线性映射，则以下属性保持不变：

1. 如果F是注射剂，那么F>是注射剂。
2. 如果F是Surjective，则F>是Injective。

证据。如果f:e→f是注射剂，那么它有一个收缩r:f→e，这样r f=ide，如果f:e→f是注射剂，那么它有一个s:f→e部分，这样f s=idf。现在，如果f:e→f是内射的，那么我们有

（r f）>=f>r>=ide，

这意味着f>是主观的，如果f是主观的，那么我们有

（f s）>=s>f>=idf，

这意味着f>是内射的。

我们还拥有以下显示Eval地图自然性的属性。提案10.6.对于任何线性图f:e→f，我们有

f>>evale=evalf f，

或者等价地，下图是通勤路线：

F>

EO\_/O\_

逃逸

/

证据。对于每一个u e和每一个\_ f，我们有

（f>>evale）（u）（）=hf>>（evale（u）），\_i

=混合气（u），f>（）i

=hf>（），用户界面

=h\_，f（u）i

=hevalf（f（u）），i

=h（evalf f）（u），i=（evalf f）（u）（）

证明F>>evale=evalf f，如权利要求所述。

如果e和f是有限维的，那么evale和evalf是同构的，所以命题10.6表明，如果我们用e的双e\_，和f的双f\_\_，那么（f>）>=f来标识e。

作为命题10.6的推论，如果dim（e）是有限的，那么我们得到ker（f>>）=evale（ker（f））。

实际上，如果e是有限维的，那么map evale:e→e\_是同构的，因此，对于某些u∈e，每个\_ e\_的形式都是\_=evale（u），map evalf:f→f\_是内射的，

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

这证明了ker（f>>）=evale（ker（f））。

下面的命题说明了正交性和换位之间的关系。

提案10.7.给定一个线性映射f:e→f，对于e的任何子空间v，我们有

f（v）0=（f>）−1（v 0）=w∈f f>（w）∈v 0。

因此，

kerf>=（imf）0和kerf=（imf>）0。

证据。我们有

hw，f（v）i=hf>（w），vi，

对于所有的V ，f（v）i=0，对于每个V，f（v）0，iff（v）0，iff（w）。

由于我们已经观察到e0=（0），在上面的恒等式中取v=e，我们得到

切口>=（imf）0.

从方程式

hw，f（v）i=hf>（w），vi，

我们推导出，v∈（imf>）0 i f f hf>（w），v i=0表示所有w∈f iff hw，f（v）i=0表示所有w∈f。假设v∈（imf>）0。如果我们选取f的基（wi）i∈i，那么我们有这样的线性形式，并且因为我们必须对所有i∈i和（wi）i∈i都有=0，我们得出f（v）=0，因此v∈kerf（这是因为hwi，f（v）i是与基向量wi关联的f（v）的系数）。相反，如果v∈kerf，那么hw，f（v）i=0代表所有w∈f，那么我们得出v∈（imf>）0。因此，v∈（imf>）0 iff v∈kerf，即：

切口=（imf>）0，

如要求。

下面的命题给出了商空间e/u的对偶（e/u）的自然解释。

提案10.8.对于向量空间e的任何子空间u，如果p:e→e/u是e/u上的标准投影，则p>是内射的，并且

im（p>）=u0=（ker（p））0.

因此，p>是（e/u）和u0之间的线性同构。

证据。由于p是主观的，根据命题10.5，映射p>是内射的。显然，u=ker（p）。观察到im（p>）由所有线性形式ψe组成，因此，对于某些\_（e/u）来说，ψ=\_p，由于ker（p）=u，我们得到u ker（ψ）。相反，对于任何线性

式ψ∈e，如果u ker（ψ），则通过e/u的ψ因子为ψ=ψp，如下图所示。

磷

E/E/U

CCCψCCCCC！γ

K

式中，ψ：e/u→k由下式给出

ψ（v）=ψ（v），v∈e，

式中，v∈e/u表示v∈e的等价类，图ψ不依赖于

在V类当量中选择的代表，因为如果v0=v，即v0−v=u∈u，则ψ（v0）=ψ（v+u）=ψ（v）+ψ（u）=ψ（v）+0=ψ（v）。因此，我们有

im（p>）=p∈（e/u）

克（ψ）

这证明了我们的结果。

命题10.8给出了对偶定理（定理10.1）第（b）部分的另一个证明，它不涉及基的存在（在无限维中）。

提案10.9.对于任何向量空间e和e的任何子空间v，我们有v 00=v。

证据。我们首先观察到v 0=v 000。这是因为，对于e的任何子空间u，我们都有u u00，所以v 0 v 000。此外，V V 00保持，任何两个子空间

m，n，e，如果m n，那么n0 n0，我们得到v 000 v 0。写v1=v 00，这样。我们想证明v1=v。

由于v v1=v 00，规范投影p1:e→e/v1因子为p1=f p，如下图所示，

磷

E/E/V型

CCPC1CCCCC！f

E/V1

式中，p:e→e/v是e/v上的正则投影，f:e/v→e/v1是p1诱导的商映射，其中f（ue/v）=p1（u）=ue/v1，对于所有u∈e（既然v v1，如果u−u0=v∈v，那么u−u0=v∈v1，那么p1（u）=p1（u0））。既然p1是主观性的，f也是主观性的，我们想证明f实际上是同构的，因此，就足以证明f是内射的。通过变换所有的映射，我们得到了交换图。

P>

e dho hhhhhhh（he/vo f>）

（e/v1），

但是根据命题10.8，映射p>：（e/v）→v 0和是同构的，因此，我们有下面的图，其中p>和都是同构的：

v 0 hdo hhph>hhhhh（he/vo f>）

（E/V1）。

因此，是同构的。我们声称这意味着f是内射的。

如果f不是内射的，那么有一些x∈e/v，这样x=06和f（x）=0，所以对于每一个θ∈（e/v1），我们有f>（）（x）=（f（x））=0。然而，存在线性形式ψ∈（e/v）使得ψ（x）=1，因此ψ=6 f>（）对于所有的ω∈（e/v1），与f>是主观的这一事实相矛盾。为了找到这样一个线性形式ψ，选取e/v中kx的任何补充w，使e/v=kx w（w是e/v中不含x的超平面），并将ψ定义为w的零和x的1。因此，f是内射的，既然我们已经知道它是外射的，它是双射的。这意味着带有v v1的正则映射f:e/v→e/v1是同构的，这意味着v=v1=v 00（否则，如果v∈v1−v，那么p1（v）=0，那么f（p（v））=p1（v）=0，但是p（v）6=0，因为v/∈v，而f不是内射的）。

下面的定理说明了f的秩与f>的秩之间的关系。

定理10.10。对于线性映射f:e→f，以下属性保持不变。（a）国际货币基金组织的双重（国际货币基金组织）与国际货币基金组织同构，即，

（国际货币基金组织）≈国际货币基金组织>。

（b）Rk（f）≤Rk（f>）。如果Rk（f）是有限的，我们得到Rk（f）=Rk（f>）。证据。（a）考虑线性图

P J E−→IMF−→F，

P.J

式中，e−→imf是e−→f诱导的推测图，imf−→f是imf的注入包含图，根据定义。为了简化符号，让i=imf。根据命题10.5，由于e−→p i是主观性的，i−p→e是内射的，并且

因为imf−→f是注射剂，所以是推测性的。既然f=j\_p，我们也有

f>=（j\_p）>=p>j>，

既然是主观的，是内射的，我们在（imf）和（f）之间有同构。

（b）我们已经注意到，定理10.1的（a）部分表明，对于每个向量空间e，dim（e）≤dim（e）。因此，dim（imf）≤dim（imf），由（a）表示，rk（f）≤rk（f>）。当dim（imf）是有限的，我们已经观察到作为定理10.1的推论，dim（imf）=dim（（imf）），因此，通过（a）部分，我们得到了rk（f）=rk（f>。

如果dim（f）是有限的，那么还有一个简单的证明（b）不使用（a）部分的结果。根据定理10.1（c）

dim（imf）+dim（imf）0）=dim（f），

根据定理5.11，dim（kerf>）+dim（imf>）=dim（f）。

此外，根据10.7号提案，我们

切口>=（imf）0，

因为f是有限维dim（f）=dim（f），所以我们推导

dim（imf）+dim（（imf）0）=dim（（imf）0）+dim（imf>），

得出dim（imf）=dim（imf>）；也就是说，rk（f）=rk（f>）。

评论：

1. 如果dim（e）是有限的，根据dan guralnik的论点，我们也可以证明rk（f）=rk（f>）如下。

我们从适用于f>：f→e的10.7号提案得知

ker（f>>）=（imf>）0，

我们从10.6号提案中得出结论：

ker（f>>）=evale（ker（f））。

因此（因为evale是同构的），dim（（im f>）0）=dim（ker（f>））=dim（ker（f））=dim（e）−dim（im f），

因为dim（imf>）+dim（（imf>）0）=dim（e），

我们得到dim（im f>）=dim（imf）。

1. 如Dan Guralnik所述，如果dim（e）是有限的，则上述结果可用于证明

imf>=（ker（f））0.

从

h f>（），u i=h\_，f（u）i

对于所有的\_ f和所有的u e，我们看到如果u ker（f），那么hf>（），ui=h\_，0i=0，这意味着f>（\_）（ker（f））0，因此，imf>（ker（f））0。反之，因为dim（e）是有限的，我们有

dim（（ker（f））0）=dim（e）−dim（ker（f））=dim（imf），

但我们刚刚证明了dim（im f>）=dim（imf），所以我们得到

dim（（ker（f））0）=dim（imf>），

由于imf>（ker（f））0，我们得到

imf>=（ker（f））0，

如要求。现在，由于（ker（f））00=ker（f），上述方程得出了另一个事实证明：

ker（f）=（imf>）0，

当e是有限维时。

三。方程式

imf>=（ker（f））0

实际上是有效的，即使当e是无限维时，如我们现在所证明的。

提案10.11.如果f:e→f是任何线性映射，则以下恒等式成立：

imf>=（ker（f））0

ker（f>）=（imf）0

imf=（ker（f>）0 ker（f）=（imf>）0.

证据。等式ker（f>）=（imf）0已经在命题10.7中得到证明。

根据对偶定理（ker（f））00=ker（f），所以从imf>=（ker（f））0我们得到ker（f）=（imf>）0。同样，（imf）00=imf，所以从ker（f>）=（imf）0我们得到imf=（ker（f>）0。因此，有待证明的是，imf>=（ker（f））0。

设P:E→E/Ker（f）为规范化假设，F:E/Ker（f）→imf为f诱导的同构，J:imf→f为包含图。那么，我们有了

F=J\_F\_P，

这意味着

F>=P>F>J>。

因为P是射血的，P>是射血的，因为J是射血的，是射血的，因为F是射血的。

双射，f>也是双射。由此得出（e/ker（f））=im（f\_j>），我们有imf>=imp>。

因为p:e→e/ker（f）是标准的投注，所以命题10.8适用于u=

克尔（F），我们得到

imf>=imp>=（ker（f））0，

如要求。

总之，方程式

imf>=（ker（f））0

适用于任何维度，它意味着

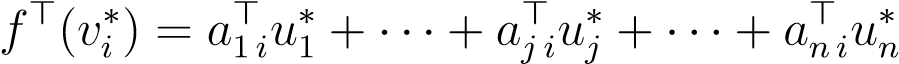
ker（f）=（imf>）0.

下面的命题显示了表示线性映射f:e→f的矩阵与表示其转置f>：f→e的矩阵之间的关系。

提案10.12。设e和f为两个向量空间，设（u1，…，un）为e的基，（v1，…，vm）为f的基。给定任意线性映射f:e→f，如果m（f）是表示f w.r.t的m×n-矩阵。基（u1，…，un）和（v1，…，vm），n×m-矩阵

m（f>）表示f>：f→e w.r.t.双碱基，是m（f）的转置m（f>。

证据。回想一下，m（f）的第i行和第j列中的条目aij是f（uj）在基（v1，…，vm）上的第i个坐标。根据vi的定义，我们得到hvi，f（uj）i=aij。m（f>）第j行和第i列中的条目是



基于（），这是公正的。自从

hvi，f（uj）i=hf>（vi），uji，

我们有a i j=a>j i，证明m（f>）=m（f>。

10.6。四个基本子空间

现在我们可以给出一个非常简短的证明，证明矩阵的秩等于其转置的秩。

提案10.13。给定一个M×N矩阵A在一个K域上，我们得到了Rk（a）=Rk（a>）。

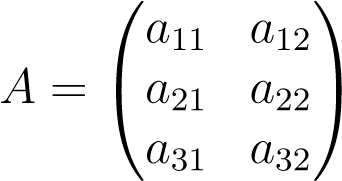
证据。矩阵A对应于线性映射f:kn→km，根据定理10.10，Rk（f）=Rk（f>）。根据命题10.12，线性映射f>对应于a>。由于Rk（a）=Rk（f），Rk（a>）=Rk（f>，我们得出Rk（a）=Rk（a>）。

因此，给定m×n矩阵a，线性无关列的最大数目等于线性无关行的最大数目。还有其他方法可以证明这个事实，不涉及对偶空间，而是行和列上的一些基本转换。

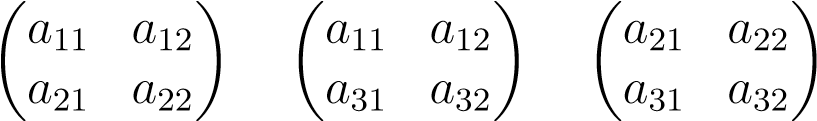
命题10.13立即给出了确定矩阵秩的下列标准：

提案10.14.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，可以得到a的可逆r×r子矩阵。

例如，3×2矩阵



具有三个2×2矩阵之一的秩2 iff



是可逆的。我们在第6章中看到，这相当于上面一个矩阵的行列式非零的事实。这不是一个非常有效的方法来寻找矩阵的秩。我们将看到有更好的方法可以使用各种分解，如LU、QR或SVD。

## 10.6四个基本子空间

给出了一个线性映射f:e→f（其中e和f是有限维），命题10.7揭示了这四个空间

国际货币基金组织，国际货币基金组织>，切口，切口>

扮演一个特殊的角色。它们通常被称为与f相关的基本子空间。这些空间以一种亲密的方式相关，因为命题10.7表明

切口=（imf>）0切口>=（imf）0，

定理10.10表明Rk（f）=Rk（f>）。

用矩阵来翻译这些关系是有指导意义的（事实上，某些线性代数书籍对此做了很大的讨论！）如果dim（e）=n和dim（f）=m，给定e的任何基（u1，…，un）和f的基（v1，…，vm），我们知道f由m×n矩阵a=（aij）表示，其中a的jth列等于f（uj）除以基（v1，…，vm）。此外，转置映射f>由n×m矩阵a>表示（相对于双碱基）。因此，四个基本空间

国际货币基金组织，国际货币基金组织>，切口，切口>

对应于

1. a的列空间，用ima或r（a）表示；这是由a的列所跨越的rm的子空间，与f的图像imf相对应。
2. a的核或空空间，用kera或n（a）表示；这是由所有向量x∈rn组成的rn的子空间，这样ax=0。
3. a的行空间，用ima>或r（a>）表示；这是用a>的行或相当于a>的列跨越的rn的子空间，它对应于f>的图像imf>。
4. 由kera>或n（a>）表示的a的左核或左零空间；这是a>的核（零空间），rm的子空间由所有向量y∈rm组成，使得a>y=0，或等价地，y>a=0。

回想一下，imf的维数r，也等于列空间的维数ima=r（a），是a（和f）的秩。然后，我们以前的一些结果可以重新表述如下：

1. a的列空间r（a）具有维度r。
2. a的空空间n（a）具有尺寸n−r。
3. 行空间r（a>）具有维度r。
4. a的左侧空空间n（a>）具有尺寸m−r。

10.6。四个基本子空间

以上陈述构成了Strang所称的线性代数基本定理，第一部分（见Strang[165]）。

这两个陈述

切口=（imf>）0

切口>=（imf）0

翻译为

1. a的空空间是a的行空间的正交。
2. a的左空空间是a的列空间的正交。

以上陈述构成了Strang所称的线性代数基本定理，第二部分（见Strang[165]）。

由于向量由列向量和线性形式由行向量表示（在e或f的基上），如果

yx=0.

然后，向量x∈rn与iff的行空间正交，x与a的每一行正交，即ax=0，相当于x属于a的空空间，同样，列向量y∈rm（表示f的对偶基上的线性形式）属于空。a>iff a>y=0，iff y>a=0，这意味着y>给出的线性形式（在f的基上）与a的列空间正交。

由于（2）等于a的列空间等于a的左零空间的正交，我们得到了形式为ax=b的方程的可解性的下列准则：

方程ax=b对所有y∈rm有一个解iff，如果a>y=0，则y>b=0。

实际上，右边的条件是b与a的左边空空间是正交的，即b属于a的列空间。

如果直接检查B是否由A列跨越，这个标准会更便宜。例如，如果我们考虑系统

x1−x2=b1 x2−x3=b2 x3−x1=b3

其矩阵形式为ax=b，如下所示：

，

我们看到矩阵A的行加起来是0。事实上，很容易让我们相信，a的左零空间的范围是y=（1,1,1），因此系统是可解的，如果y>b=0，即

b1+b2+b3=0。

请注意，上述标准也可作如下负面说明：

方程ax=b没有解，如果有y∈rm，则a>y=0，y>b=0.6。

由于a>y=0，当y>a=0时，我们可以将y>视为表示线性形式的行向量，y>a=0断言线性形式y>在a1列上消失，…，a的a列上消失，但不在b上消失。由于线性形式y>定义了方程y>z=0的超平面h（用z∈rm），几何上方程ax=b没有解，如果有一个超平面h，包含a1，…，a，不包含b。

## 10.7总结

本章的主要概念和结果如下：

* 双空间E和线性形式（covector）。招标人E。
* 双线性配对H−、−I:E×E→K（规范配对）。
* 在v:evalv:e→k时的评估。
* 地图评估：E→E。
* e的子空间v与e的子空间u之间的正交性；正交v 0和正交u0。
* 坐标形式。
* 对偶定理（定理10.1）。
* 基础的双重基础。
* 当dim（e）有限时，同构evale:e→e。
* 两个向量空间之间的配对；非退化配对；命题10.3。
* 超平面和线性形式。
* 线性映射f:e→f的转置f>：f→e。
* 基本特征：

kerf>=（imf）0和kerf=（imf>）0

（提案10.7）。

10.7。总结

* 如果f是有限维，那么

Rk（f）=Rk（f>）。

（定理10.10）。

* 转置映射f>的矩阵等于映射f的矩阵的转置（命题10.12）。
* 对于任何M×N矩阵A，Rk（a）=Rk（a>）。•根据最大可逆子矩阵描述矩阵的秩（命题10.14）。
* 四个基本子空间：

国际货币基金组织，国际货币基金组织>，切口，切口>。

* （矩阵的）列空间、空空间、行空间和左空空间。
* 形式为ax=b的方程的左零空间可解性准则。

350第10章。二元空间，二元性

第十一章

# 欧几里得空间

最美丽的人。

-赫尔曼·明可夫斯基

## 11.1内部产品、欧几里得空间

到目前为止，向量空间的框架允许我们处理向量的比例和线性组合，但是没有办法表达角度的概念或讨论向量的正交性。欧几里得结构允许我们处理度量概念，如角度、正交性和长度（或距离）。

本章介绍欧几里得几何的基本内容。第12章研究了欧几里得几何的深层次。我们的主要目标之一是给出保持欧几里得结构、旋转和反射的变换的基本性质，因为它们在实践中起着重要作用。欧几里得几何学是研究某些仿射映射（称为刚性运动）下不变性质的学科。刚性运动是保持点之间距离的贴图。

我们首先定义内部积和欧几里得空间。给出了柯西-施瓦兹不等式和明可夫斯基不等式。我们定义了向量和子空间、正交基和正交基的正交性。我们证明了每个有限维欧氏空间都有正交基。第一个证明使用对偶性，第二个证明使用格拉姆-施密特正交化过程。可逆矩阵的QR分解是格拉姆-施密特程序的一个应用。简单地定义和研究了线性等轴测（也称为正交变换）。最后，我们用一个简短的部分来总结欧几里得几何的一些应用。最重要的应用之一，最小二乘法，将在第21章中讨论。

关于欧几里得几何学的更详细的处理，请参见Berger[11，12]、Snapper和Troyer[157]或任何其他几何学书籍，如Pedoe[132]、Coxeter[44]、Fresnel[66]，

Tisseron[170]或Cagnac、Ramis和Commeau[32]。严肃的读者应该咨询埃米尔。

三百五十一

Artin的著名著作[6]对正交群以及其他几何群进行了深入的研究。仍然值得参考一些古老的经典著作，如Hadamard[84，85]和Rouch'e和De Combourse[135]。[84]的第一版于1898年出版，最终于1947年出版第十三版！在本章中，假设所有向量空间都定义在实数的字段r上，除非另有规定（在少数情况下，定义在复数c上）。

首先，我们在向量空间上定义一个欧几里得结构。从技术上讲，矢量空间E上的欧几里得结构是由满足某些额外性质的矢量空间上的对称双线性形式提供的。回想一下双线性形式，如果对于每个u∈e，u=06，则定义为：e×e→r；如果对于每个u∈e，则定义为：θ（u，u）=06；如果对于每个u∈e，则定义为正。

定义11.1.欧几里得空间是一个实向量空间e，它具有对称双线性形式，即：e×e→r，是正定的。更明确地说，\_：e×e→r满足

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

实数（u，v）也被称为u和v的内积（或标量积）。我们还定义了与之相关联的二次型为函数Φ：e→r+，这样

Φ（u）＝（u，u）、

对于所有的u∈e。

既然π是双线性的，我们就有了π（0,0）=0，既然它是正定的，我们就有了更强有力的事实：

⑨（u，u）=0 iff u=0，

也就是说，当u=0时，Φ（u）=0。

给定向量空间e上的内积，我们也用

u·v或hu、vi或（u v）

以及Kuk的pΦ（u）。

例11.1。欧几里得空间的标准例子是Rn，在内积下定义如下：

（x1，…，xn）·（y1，…，yn）=x1y1+x2y2+····+xnyn.

这个欧几里得空间用en表示。

还有其他的例子。

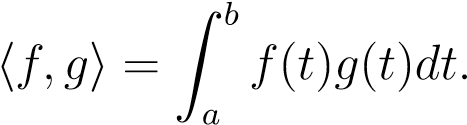
例11.2。例如，让e是维度2的向量空间，让（e1，e2）是e的基础。如果a>0和b2−ac<0，双线性形式定义如下：

⑨（x1e1+y2e2，x2e1+y2e2）=axix2+b（x1y2+x2y1）+cy1y2

在E上生成一个欧几里德结构。在这种情况下，

Φ（Xe1+Ye2）=ax2+2bxy+cy2。

例11.3。让c[a，b]表示一组连续函数f[a，b]→r。很容易检查c[a，b]是一个无限维的向量空间。给定任意两个函数f，g∈c[a，b]，让



我们把它作为一个简单的练习，即H−，−I实际上是C[A，B]上的一个内积。在A=−π和B=π（或A=0和B=2π，这基本上没有区别）的情况下，应该计算

HSINPX、SINQXI、HSINPX、COSQXI和HCOSPX、COSQXI，

对于所有自然数p，q≥1。这些计算的结果使傅立叶分析成为可能！

例11.4。设e=mn（r）为实n×n矩阵的向量空间。如果我们把一个矩阵a∈mn（r）看作是一个“长”列向量，把它的列连在一起，我们就可以把两个矩阵a，b∈mn（r）的内积定义为

，

可以方便地写为

ha，bi=tr（a>b）=tr（b>a）。

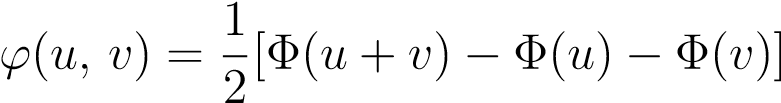
因为这可以看作是RN2上的欧几里得积，所以它是mn（r）上的内积。相应的规范

kakf=ptr（a>a）

是弗罗贝尼乌斯规范（见第8.2节）。

让我们观察一下，可从Φ中恢复。

提案11.1.我们有



对于所有的u，v∈e，我们说，a是Φ的极性形式。

证据。根据双线性和对称性，我们有

Φ（u+v）=（u+v，u+v）

=\_（u，u+v）+\_（v，u+v）

=\_（u，u）+2（u，v）+\_（v，v）=Φ（u）+2（u，v）+Φ（v）。

如果e是有限维，如果\_：e×e→r是e上的双线性形式，给定任何基础

（e1，…

.

如果我们让g为矩阵g=（（ei，ej）），如果x和y是与（x1，…，xn）和（y1，…，yn）相关联的列向量，那么我们可以写

⑨（x，y）=x>gy=y>g>x。

注意，我们犯了一个符号滥用的错误，因为它是e中的向量，但是与（x1，…，xn）相关联的列向量属于rn。为了避免这种轻微的滥用，我们可以用x表示与（x1，…，xn）相关联的列向量（对于与（y1，…，yn）相关联的列向量，也可以用y表示），在wich情况下，“正确”的表示是

⑨（x，y）=x>gy。

然而，考虑到e和rn之间的同构，为了尽可能简单地保持符号，我们将使用x和y而不是x和y。

另外，观察到\_为对称iff g=g>，且\_为正定iff，矩阵g为正定，即，

x>gx>0表示所有x∈rn，x=06。

与内积相关的矩阵G称为内积相对于基（e1，…，en）的G矩阵。

相反，如果a是对称正定n×n矩阵，则很容易检查双线性形式

hx，yi=x>ay

是内部产品。如果我们把基从基（e1，…，en）改为基（f1，…，fn），如果基矩阵的变化是p（其中p的jth列由基（e1，…，en）上的fj坐标组成），那么对于基（f1，…，fn）上的x0和y0坐标，我们有

x>gy=x0>p>gpy0，

因此，我们内部产品的基矩阵（f1，…，fn）是p>gp。我们将这些事实概括为以下命题。

提案11.2.设e为有限维向量空间，设（e1，…，en）为e的基础。

1. 对于任何内积H−、−i on e，如果g=（hei，eji）是内积H−、−i w.r.t的克矩阵。基（e1，…，en），则g是对称正定的。
2. 对于基矩阵p的任何变化，相对于新基，h−、−i的g矩阵是p>gp。
3. 如果a是任意n×n对称正定矩阵，则

hx，yi=x>ay

是E上的内部产品。

稍后我们将看到一个对称矩阵是正定的，如果它的特征值都是正的。

内积的一个非常重要的性质是，图U 7→PΦ（U）是一个范数。

提案11.3.设e为内积\_的欧几里得空间，设Φ为相应的二次型。对于所有的u，v∈e，我们有柯西-施瓦兹不等式。

⑨（u，v）2≤Φ（u）Φ（v）、

当f u和v相等时，它们是线性相关的。

我们还有明可夫斯基不平等

磷磷

Φ（U+V）≤Φ（U）+Φ（V）、

等式保持iff u和v是线性相关的，此外，如果u=06和v=06，则对于某些λ>0，u=λv。

证据。对于任何向量u，v∈e，我们定义函数t:r→r，这样

t（λ）=Φ（u+λv）

对于所有的λ∈r，利用双线性和对称性，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 |

既然ω是正定的，那么Φ是非负的，因此t（λ）≥0代表所有的λ∈r。如果Φ（v）=0，那么v=0，我们也有（u，v）=0。在这种情况下，柯西-施瓦兹不等式是微不足道的，v=0和u是线性相关的。

现在假设Φ（v）>0。由于t（λ）≥0，二次方程

λ2Φ（v）+2λ（u，v）+Φ（u）=0

不能有明显的实根，这意味着它的判别式

∆=4（\_（u，v）2−Φ（u）Φ（v））

是零还是负，这正是柯西-施瓦兹不等式。

⑨（u，v）2≤Φ（u）Φ（v）。

现在让我们考虑一下我们有平等的情况

⑨（u，v）2=Φ（u）Φ（v）。

有两种情况。如果Φ（v）=0，则v=0，u和v是线性相关的。如果Φ（v）=06，则上述二次方程有一个双根λ0，我们得到Φ（u+λ0v）=0。由于\_为正定，因此Φ（u+λ0v）=0表示u+λ0v=0，这表明u和v是线性相关的。相反，当u和v是线性相关时，很容易检查我们是否具有相等性。明可夫斯基不平等

磷磷

Φ（U+V）≤Φ（U）+Φ（V）

等于

Φ（U+V）≤Φ（U）+Φ（V）+2PΦ（U）Φ（V）。

但是，我们已经证明

2Ф（u，v）=Φ（u+v）−Φ（u）−Φ（v）、

所以上面的不等式等于

⑨（u，v）≤pΦ（u）Φ（v）、

式中，当ω（u，v）≤0时，这是微不足道的，当ω（u，v）≥0时，从柯西-施瓦兹不等式得出。因此，明可夫斯基不平等是成立的。最后假设u=06，v=06，那么

磷磷

Φ（U+V）=Φ（U）+Φ（V）。

在这种情况下，我们

⑨（u，v）=pΦ（u）Φ（v）、

通过对柯西-施瓦兹不等式的讨论，我们知道，当u和v是线性相关的时，等式成立。当u或v为空时，minkowski不等式是一个等式。否则，如果u=06，v=06，那么对于某些λ=06，u=λv，并且由于

⑨（u，v）＝λ⑨（v，v）=pΦ（u）Φ（v）、

根据正性，我们必须有λ>0。

注意，柯西-施瓦兹不等式也可以写成

|⑨（U，V）≤PΦ（U）PΦ（V）。

注：很容易证明柯西-施瓦兹不等式和Minkowski不等式仍然适用于一个正的对称双线性形式，但不一定是确定的（即，所有u，v∈e，u，v）≥0）。然而，当等式成立时，u和v不需要线性依赖。

明可夫斯基不平等

磷磷

Φ（U+V）≤Φ（U）+Φ（V）

结果表明，映射u 7→pΦ（u）满足定义8.1的凸不等式（又称三角不等式）、条件（n3），且由于π为双线性正定，因此它也满足定义8.1的条件（n1）和（n2），因此它是e上的范数，由π引起的范数。被称为\_诱导的欧几里得范数。柯西-施瓦兹不等式可以写成

| U·V≤Kukkkvk，

以及Minkowski不等式

ku+vk≤kuk+kvk。

如果u和v是非零向量，那么柯西-施瓦兹不等式意味着

.

然后有一个独特的θ∈[0，π]，这样

.

我们有u=v iffθ=0和u=−v iffθ=π。对于0<θ<π，矢量su和uv，使得v是线性独立的，平面的方向由

θ是u和v之间的角度。关于方向的精确概念，请参见问题11.8。如果u是单位向量（也就是说kuk=1），那么向量

（kvkcosθ）u=（u·v）u=（v·u）u

称为V向U所跨越的空间的正交投影。

注：人们可能会想，向量空间上的每一个范数是否都是由某个欧几里得内积引起的。一般来说，这是错误的，但值得注意的是，有一个简单的必要和充分条件，即范数必须满足平行四边形定律：

ku+vk2+ku−vk2=2（kuk2+kvk2）。

见图11.1。

u

v

u

+

v

u

-

v

图11.1：平行四边形定律表明，由向量u和v确定的平行四边形对角线长度之和等于所有边的和。如果H−，−I是内积，那么我们有

ku+vk22=kuk22+kvk22+2hu，vi

ku−vk=kuk+kvk−2hu，vi，

通过加减这些恒等式，我们得到了平行四边形定律和方程

，

这使我们能够从标准中恢复H−I。

相反，如果k k是一个满足平行四边形定律的范数，并且它来自一个内积，那么这个内积必须由

.

我们需要证明上述形式确实是对称的双线性的。

可变对称性之所以存在是因为。根据平行四边形定律，我们有k u−v k=k−（u−v）k=kv−uk。让我们证明

2（KX+ZK2+KYK2）=KX+Y+ZK2+KX−Y+ZK2

会产生

KX+Y+ZK22=2（KX+ZK22+KYK22）−KX−Y+ZK22

KX+Y+ZK=2（KY+ZK+KXK）−KY−X+ZK，

其中第二个公式是通过交换x和y得到的，然后将这些公式相加，我们得到

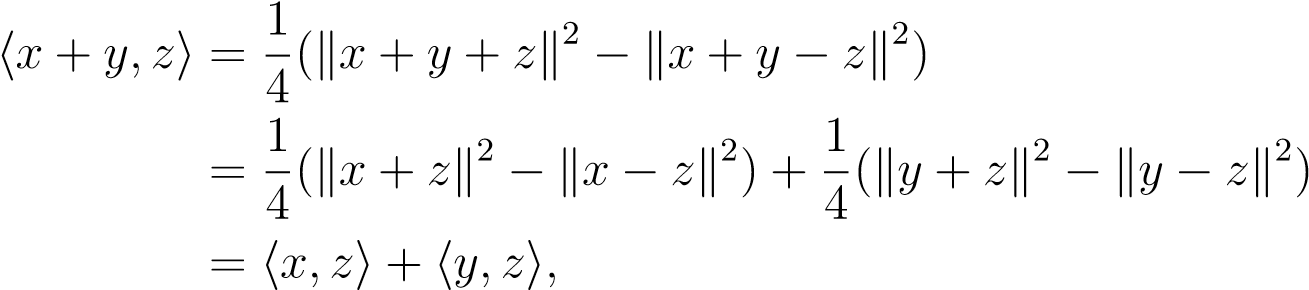
.

用上式中的-z替换z，我们得到

，

sincex−yk−xx+z）k=

钾

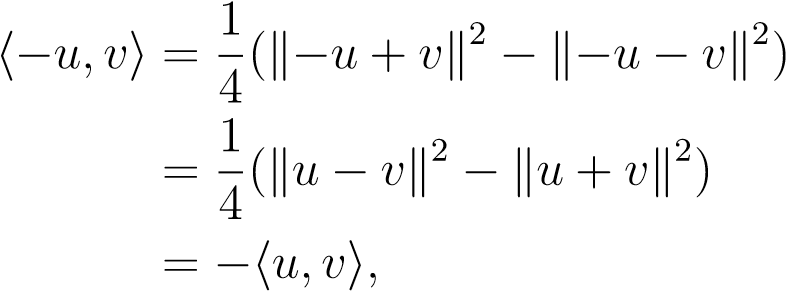


根据需要。证明

hλx，yi=λhx，yi代表所有λ∈r

有点棘手。其策略是先证明λ∈z的恒等式，然后将其推广到q，再通过连续性将其推广到r。

自从



该财产持有λ=−1。通过线性和归纳，对于任意n∈n，n≥1，写n=n−1+1，我们得到

hλx，yi=λhx，yi代表所有λ∈n，

由于上面也适用于λ=−1，所以它适用于所有的λ∈z。对于λ=p/q，p，q∈z和q 6=0，我们有qh（p/q）u，vi=hpu，vi=phu，vi，

表示h（p/q）u，vi=（p/q）hu，vi，

因此，hλx，yi=λhx，yi对于所有的λ∈q。为了完成证明，我们使用了一个范数是一个连续的映射x 7→kxk的事实。然后，在r−0上定义的连续函数与hu一致，在q−0上定义vi，因此它等于hu，在r−0上定义vi。λ=0的情况是微不足道的，所以我们结束了。

我们现在定义正交性。

## 11.2欧几里得空间的正交性和对偶性

向量空间上的内积可以定义正交性的概念。非零对正交向量族必须是线性无关的。它们被称为正交族。在有限维的向量空间中，总是可以找到正交基。这在理论上和实践上都非常有用。实际上，在正交的基础上，找到向量的坐标是非常便宜的：它需要一个内积。傅立叶级数充分利用了这个事实。当e有有限维时，我们证明了e上的内积在e与其对偶空间e之间产生了自然同构。这允许我们以一种内在的方式（即独立于基）定义线性映射的伴随。也可以对任何基础进行正交规范化（当然，当维度是有限的时）。我们给出了两个证明，一个使用对偶，另一个使用格拉姆-施密特正交化程序更具建设性。

定义11.2.在欧几里得空间E中，任意两个向量u，v∈e是正交的，或者是垂直的，如果u·v=0。给定e中向量的族（ui）i∈i，我们认为（ui）i∈i是正交的，如果ui·uj=0代表所有i，j∈i，其中i=6 j。我们说，如果ui·uj=0代表所有i，j∈i，其中i=6 j，而kuik=ui·ui=1，代表所有i∈i。

对于e的任何子集f，集合

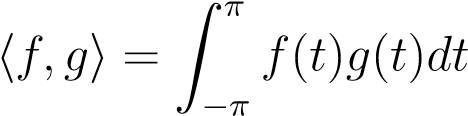
f=v∈e u·v=0，对于所有u∈f，

在所有与f中所有向量正交的向量中，称为f的正交补。

由于内积是正定的，所以观察到对于任何向量u∈e，我们有

u·v=0表示所有v∈e iff u=0。

立即证实f的正交补码f是e的子空间。例11.5。回到示例11.3和内部产品



在向量空间c[−π，π]上，很容易检查

1，

0，如果p=6 q，p，q≥1，

1，

PX，QX

如果p=6 q，p，q≥0，

和

Xinpx，cosqxi=0，

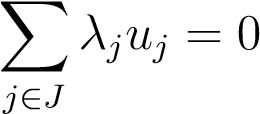
对于所有p≥1和q≥0，当然，

因此，家族（sinpx）p≥1（cosqx）q≥0是正交的。不正态，√√

但如果我们把每个三角函数除以π，1除以2π，就会变成这样。

提案11.4.给定欧几里德空间e，对于e中非零向量的任何族（ui）i∈i，如果（ui）i∈i是正交的，那么它是线性无关的。

证据。假设存在线性相关性



对于i的某个λj∈r和某个有限子集j，通过对任意i∈j取带ui的内积，利用内积的双线性度和当i=6j时ui·uj=0的事实，我们得到

！

0=ui·0=ui·xλjuj

J·J·J

=xλj（ui·uj）=λi（ui·ui），

J·J·J

所以

λi（ui·ui）=0，对于所有i∈j，

由于ui=06，且内积为正定，ui·ui=06，因此我们得到

λi=0，对于所有i∈j，

说明了族（ui）i∈i是线性独立的。

我们将下面的简单结果留作练习。

提案11.5。给定欧几里得空间e，任意两个向量u，v∈e是正交iff

ku+vk2=kuk2+kvk2。

几何解释见图11.2。

u

+

v

u

v

图11.2：直角三角形两边的长度之和等于斜边的长度，即勾股定理。

正交基的一个最有用的特征是，它们提供了一种计算任意基向量上向量坐标的非常简单的方法。实际上，假设（e1，…，em）是一个正态基。对于任何向量

x=x1e1+·····+xmem，

如果我们计算内积x·ei，我们得到

1. ·EI＝X1E1·EI+··+ XIEI·EI+···XMEM·EI＝XI，

自从

，

正态族的特征。因此，

1. =x·ei，

也就是说，xiei=（x·ei）ei是x对基向量ei生成的子空间的正交投影。见图11.3。如果基是正交的，但不一定

e

i

x

x e

i

i

*Θ*

图11.3：红色矢量x在黑色基矢量ei上的正交投影是褐红色矢量xiei。观察x·ei=kxkcosθ。

正态，那么

.

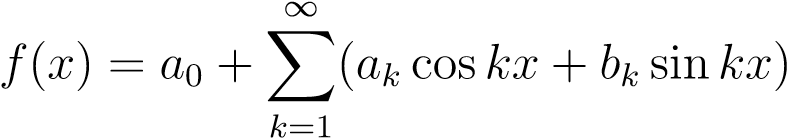
即使对于无穷大正交（或正交）基（ei）i∈i，所有这些都是正确的。

但是，记住每个向量x都表示为线性组合

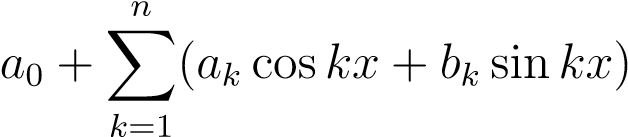
x=xxiei

我爱我

其中标量族（Xi）i i具有有限的支持，这意味着对于所有i i j j，Xi＝0，其中j是有限集。因此，即使族（sinpx）p≥1（cosqx）q≥0是正交的（它不是正交的，但如果我们将每个三角函数除以√π，1除以√2π，就会变成正交的，我们不会因为它看起来杂乱！），函数f∈c0[−π，π]可以写成傅立叶级数



并不意味着（sinpx）p≥1（cosqx）q≥0是函数向量空间的基础，因为一般来说，族（ak）和（bk）没有有限的支持！为了使这种无限线性组合有意义，有必要证明部分和



当n变为无穷大时，级数的收敛到极限。这需要空间上的拓扑结构。

有限维欧几里德空间的一个非常重要的性质是内积在向量空间E与其对偶E之间产生一个规范的双射（即不依赖于基的选择）。原因是内部产品·：e×e→r定义了一个非退化配对，定义见10.4。实际上，如果u·v=0表示所有v∈e，则u=0；同样，如果u·v=0表示所有u∈e，则v=0（因为内积是正定对称的）。根据命题10.3，E和E之间存在一个规范同构。我们认为，如果我们显式地展示这个映射，并谴责它是同构的话，读者会感激的。

从E到E的映射定义如下。

定义11.3.对于任何向量u∈e，设u:e→r为定义为

u（v）=u·v，对于所有v∈e。

由于内积是双线性的，所以图\_u是e\_中的线性形式。因此，我们有一个地图[：e→e，定义如下：

[（u）=u.

定理11.6。给定欧几里得空间e，地图[：e→e定义如下：

[（u）=u

是线性的和内射的。当e也是有限维时，映射[：e→e是一个正则同构。

证据。[：e→e是一个线性映射，它直接从内积是双线性这一事实出发。如果u=v，则所有w∈e的u（w）=v（w），根据u的定义，这意味着所有w∈e的u·w=v·w，通过双线性度，这相当于

（V−U）·W=0

对于所有的w∈e，这意味着u=v，因为内积是正定的。因此，[：e→e是内射的。最后，当e是有限维n时，我们知道e也是维n，那么[：e→e是双射的。

同构的逆[：e→e用]：e→e表示。

由于定理11.6，我们有以下推论。

推论11.7。如果e是有限维的欧几里德空间，则每一个线性形式f∈e都对应一个唯一的u∈e，这样

f（v）=u·v，每v∈e。

特别地，如果f不是零形式，那么f的核心，即超平面h，就是一组与u正交的向量。

评论：

1. “音乐地图”[：e→e在e具有无限维度时不是主观的。通过将我们的注意力限制在连续线性映射上，并假设向量空间e是希尔伯特空间（即，e是一个完全赋范向量空间w.r.t.欧几里得范数），可以挽救结果。这就是著名的“小”里兹定理（或里兹表示定理）。
2. 定理11.6仍然成立，如果E上的内积被非退化对称双线性形式\_替换。我们说对称双线性形式，如果每一个u∈e，那么它是非退化的。

如果所有v∈e中的（u，v）=0，则u=0。

例如，R4上的对称双线性形式（洛伦兹形式）定义如下：

⑨（（x1，x2，x3，x4），（y1，y2，y3，y4））=x1y1+x2y2+x3y3−x4y4

是不退化的。然而，也存在非零向量u∈r4，使得（u，u）=0，这在欧几里得空间中是不可能的。这种矢量称为各向同性。

例11.6。考虑RN和它通常的欧几里得内积。给定任意可微函数f:u→r，其中u是Rn的开子集，根据定义，对于任意x∈u，f在x处的总导数dfx是定义的线性形式，因此对于所有u=（u1，…，un）∈Rn，

.

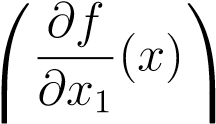
唯一向量v∈rn，这样

v·u=dfx（u）表示所有u∈rn

是f的雅可比矩阵在x处的转置，1×n矩阵

.

这是f在x处的梯度梯度（f）x，由



梯度（.

例11.7。给定任意两个向量u，v∈r3，设c（u，v）为

c（u，v）（w）=det（u，v，w），对于所有w∈r3。

自从

，

我们看到唯一向量z∈r3，这样

z·w=c（u，v）（w）=det（u，v，w），对于所有w∈r3

是向量

.

-

这只是u和v的叉积u×v。由于det（u，v，u）=det（u，v，v）=0，我们发现u×v与u和v都是正交的。上面允许我们将叉积推广到rn。给定任意n-1向量u1，…，un-1∈rn，交叉积u1×·····×un-1是rn中的唯一向量，这样

（u1×·····×un-1）·w=所有w∈rn的det（u1，…，un-1，w）。

11.3。线性映射的伴随

例11.8。考虑内积为ha，bi=tr（a>b）的实n×n矩阵的向量空间mn（r）。设为：mn（r）→r为

，

其中a=（aij）。立即证明S是线性形式。很容易检查唯一矩阵z

hz，ai=s（a），对于所有a∈mn（r）

是矩阵z=一（n，n），其条目均等于1。

## 11.3线性图的伴随

同构的存在对伴随图的存在至关重要。伴随映射的重要性源于物理问题中产生的线性映射往往是自伴随的，这意味着f=f。此外，自伴映射可以在特征向量的正交基上对角化。这是解决力学和工程中许多问题的关键（见Strang[164]）。

设e为有限维n的欧几里得空间，设f:e→e为线性映射。对于每一个u∈e，图v 7→u·f（v）

显然是e中的线性形式，根据定理11.6，e中有一个用f（u）表示的唯一向量，使得f（u）·v=u·f（v），

对于每一个v e，下面的简单命题表明映射f是线性的。

提案11.8.给定一个有限维的欧几里得空间e，对于每一个线性映射f:e→e，都有一个唯一的线性映射f：e→e，这样

f（u）·v=u·f（v），对于所有u，v∈e。

证据。给定u1，u2∈e，由于内积是双线性的，我们得到

（u1+u2）·f（v）=u1·f（v）+u2·f（v），

对于所有v∈e，和

（f（u1）+f（u2））·v=f（u1）·v+f（u2）·v，

对于所有的v∈e，由于假设，

f（u1）·v=u1·f（v）和f（u2）·v=u2·f（v），

对于所有的v∈e，因此我们得到

（f（u1）+f（u2））·v=（u1+u2）·f（v）=f（u1+u2）·v，

对于所有的v∈e，因为[是双目标的，这意味着

f（u1+u2）=f（u1）+f（u2）。

同样地，

（λu）·f（v）=λ（u·f（v）），

对于所有v∈e，和

（λf（u））·v=λ（f（u）·v），

对于所有的v∈e，由于假设，

F（U）·V=U·F（V）

对于所有的v∈e，我们得到

（λf（u））·v=λ（u·f（v））=（λu）·f（v）=f（λu）·v

对于所有的v∈e，因为[是双目标的，这意味着

f（λu）=λf（u）。

因此，F实际上是一个线性映射，它是独一无二的，因为[是一个双射体]。

定义11.4.给定一个有限维的欧几里得空间e，对于每一个线性映射f:e→e，唯一的线性映射f：e→e，这样

f（u）·v=u·f（v），对于所有u，v∈e

由命题11.8给出的称为f（w.r.t.与内积）的伴随。线性映射f:e→e这样f=f被称为自伴映射。

自伴线性映射具有实特征值，正态基是由特征向量产生的，因而起着非常重要的作用。此外，许多物理问题导致自伴线性映射（以对称矩阵的形式）。

注：如果e上的内积被非简并对称双线性形式\_替换，则命题11.8仍然成立。

11.3。线性映射的伴随

线性映射，使f−1=f，或等效

F F=F F=ID，

也扮演着重要的角色。它们是线性等轴测或等轴测。旋转是一种特殊的等角图。另一类重要的线性映射是满足该性质的线性映射。

F F=F F，

称为法向线性映射。稍后我们将看到，法向映射总是可以在特征向量的正交基上对角化，但这需要使用厄米特内积（超过c）。

两个欧几里得空间f上的内积表示为by e和f，其中h−、−i e上的内积表示为h−f，它是−i1，立即验证了命题11.8的证明可以被调整，以表明其中2，给定任何线性映射f:e→是唯一的线性映射f：f→e这样

hf（u），vi2=hu，f（v）i1

对于所有u∈e和所有v∈f，线性映射f也被称为f的伴随。

从伴随映射的定义可以看出以下特性：（1）对于任何线性映射f:e→f，我们有

F=F.

1. 对于任意两个线性映射f，g:e→f和任意标量λ∈r：
   1. +g）=f+g（λf）=λf。
2. 如果e、f、g、−i3和ife是具有各自内部产物的欧几里德空间f:e→f和g:f→g是两个线性映射，则h−、−i1、h−、−i2和h−
   1. \_f）=f g。

注：考虑到e的任何基和f的任何基，可以根据f的矩阵和定义内积的革兰氏矩阵来描述f的伴随f的矩阵；见问题11.5。我们将在11.14（2）号提案中对正态基这样做。而且，由于内积是对称的，f的伴随f\_也具有f（u）·v=u·f（v）的特征。

对于所有u，v∈e。

## 11.4正交基底的存在和建造

我们也可以利用定理11.6来证明有限维的欧几里德空间具有正交基。

提案11.9.给定任意非平凡欧几里得空间e的有限维n≥1，e有一个正交基（u1，…，un）。

证据。我们在n上进行归纳，当n=1时，取任意一个非零向量v∈e，它存在于我们假定e非平凡之后，并让

.

如果n≥2，再取任意非零向量v∈e，设

.

考虑与U1相关的线性形式U1。由于u1=06，根据定理11.6，线性形式u1是非零的，其核是超平面h。由于u1（w）=0 iff u1·w=0，超平面h是u1的正交补数。此外，由于U1=06且内积为正定，U1·U1=06，因此，U1∈/h，这意味着E=H Ru1。然而，由于e是有限维n，超平面h的维数为n-1，根据归纳假设，我们可以找到h的正交基（u2，…，un）。现在，由于h和一维空间ru1是正交的，e=h ru1，很明显（u1，…，un）是正交基。S代表E。

作为命题11.9的结果，给定任意一个欧几里得空间的有限维n，如果（e1，…，en）是e的正交基，那么对于任意两个向量u=u1e1+·········+unen和v=v1e1+········+vnen，内积u·v表示为

，

和诺姆库克一样

.

欧几里得空间总是有一个正交基，这意味着任何一个G矩阵都可以写成

g=q>q，

对于一些可逆矩阵q，我们知道在基矩阵的变化中，一个g矩阵变成g0=p>gp。如果与g0对应的基是正交的，则g0=i，因此g=（p−1）>p−1。

有一种更具建设性的方法来证明11.9号提案，使用一种称为格拉姆-施密特正交化程序的程序。除此之外，Gram–Schmidt正交化程序生成矩阵的QR分解，这是数值方法中的一个重要工具。

提案11.10。对于有限维n≥1的任意非平凡欧几里德空间e，从e的任意基（e1，…，en），我们可以构造e的正交基（u1，…，un），其性质是：对于每k，1≤k≤n，族（e1，…，ek）和（u1，…，uk）生成相同的子空间。

证据。我们在n上归纳n=1，让

.

对于n≥2，我们也允许

，

假设（u1，…，uk）是生成与（e1，…，ek）相同子空间的正交系统，对于1≤k<n的每个k，我们注意到矢量

钾

U0K+1=EK+1−X（EK+1·UI）用户界面

i＝1

不为空，因为（u1，…，uk）和（e1，…，ek）生成相同的子空间，（e1，…，ek+1）将线性相关，这是荒谬的，因为（e1，…，en）是基础。因此，向量的范数为非零，我们使用向量UK和U0K的以下构造：

，

对于感应步骤

，

其中1≤k≤n−1。很明显，kuk+1k=1，由于（u1，…，uk）是一个正态系统，我们有

，

对于1≤i≤k的所有i，这表明族（u1，…，uk+1）是正交的，由于（u1，…，uk）和（e1，…，ek）生成相同的子空间，从uk+1的定义可以清楚地看出（u1，…，uk+1）和（e1，…，ek+1）生成相同的子空间。这就完成了归纳步骤和命题的证明。

请注意，U0K+1是通过从EK+1减去EK+1本身在已经计算的正交向量U1，…，UK上的投影得到的。然后将U0K+1归一化。

例11.9。对于此过程的特定示例，让e=r3与标准

欧几里得标准。以此为基础

然后

和

.

这意味着

，

还有那个

.

为了完成正交基，归一化得到

.

图11.4给出了该示例的说明。

评论：

1. 现在可以很容易地获得QR分解，但我们将此推迟到第11.6节。
2. 命题11.10的证明也适用于e的可数无穷大基，产生可数无穷大正交正规基。

e

2

u

u

1

‘

2

u

1

direction

u

2

direction

e

3

u

3

‘

图11.4：上图显示蓝色的构造与e2在u1上的正交投影垂直，下图显示绿色的构造与u1和u2确定的平面垂直。

也应该说，我们提出的格拉姆-施密特正交化程序在数值上不是很稳定，而是应该使用改进的格拉姆-施密特方法。要计算，而不是在U1上投影EK+1，…，在一个步骤，英国，它是更好的执行K预测。我们计算如下：

UK1+1=EK+1−（EK+1·U1）U1，UKI+1+1=UKI+1−（UKI+1·UI+1）UI+1，

其中1≤i≤k−1。很容易看出U0K+1=UK+1。

例11.10。让我们将修改后的Gram–Schmidt方法应用于示例11.9的（e1、e2、e3）基础。唯一的变化是计算。对于修改后的Gram–Schmidt程序，我们首先计算

1 1 1

U31=E3−（E3·U1）U1=1−2/3 1=1/3 1。

0 1−2

然后

1 1 1

u32=u31−（u31·u2）u2=1/3 1+1/6−2=1/2 0，

−2 1−1

u

1

direction

u

2

direction

e

3

u

3

1

u

1

direction

u

2

direction

u

1

3

u

3

2

图11.5：上图显示蓝色的构造与E3在U1上的正交投影垂直，下图显示天空蓝色的构造与U2上的正交投影垂直。

观察一下。见图11.5。

下面的matlab程序实现了修改后的gram-schmidt过程。

函数q=gramschmidt4（e）n=size（e，1）；对于i=1:n q（：，i）=e（：，i）；对于j=1:i-1 r=q（：，j）'\*q（：，i）；q（：，i）=q（：，i）-r\*q（：，j）；

结束

r=sqrt（q（：，i）'\*q（：，i））；q（：，i）=q（：，i）/r；结束

如果我们将上述函数应用于矩阵

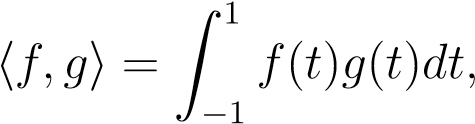
，

输出是矩阵

，

与示例11.9的结果相匹配。

例11.11。如果我们考虑多项式和内积



将格拉姆-施密特正交化程序应用于多项式

1，x，x2，…，xn，…，

作为一个实系数变量多项式的基础，我们得到了一个与勒让德多项式有关的正交多项式Qn（x）。

勒让德多项式pn（x）具有许多优良的性质。它们是正交的，但它们的范数并不总是1。勒让德多项式pn（x）的定义如下。设fn为函数fn（x）=（x2-1）n，

我们定义Pn（x）如下：

p0（x）=1，和，

式中，是fn的第n个导数。

它们也可以归纳地定义为：

.

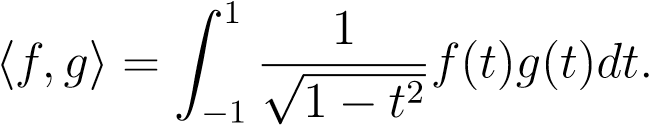
以下是pn（x）的显式求和：

.

多项式qn与勒让德多项式pn有关，如下：

.

例11.12。考虑具有对称双线性形式的[-1,1]上的多项式



我们把它作为一个练习来证明上面定义了一个内部产品。可以证明，多项式tn（x）由

tn（x）=cos（麻醉剂），n≥0，

（等价地，当x=cosθ时，我们得到tn（cosθ）=cos（nθ））与上述内积正交。这些多项式是切比雪夫多项式。它们的范数不等于1。相反，我们有

，

利用恒等式（cosθ+isinθ）n=cosnθ+isinθ和二项式，我们得到了tn（x）的以下表达式：

.

切比雪夫多项式的归纳定义如下：

t0（x）=1

T1（x）=x

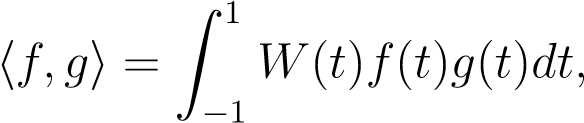
tn+1（x）=2xtn（x）−tn−1（x），n≥1.

利用这些递推方程，我们可以证明

.

多项式tn在区间[-1,1]中有n个不同的根。切比雪夫多项式在近似理论中起着重要作用。它们被用作SUP范数（∞-范数）下连续函数的最佳多项式近似。

后两个例子的内积是形式内积的特例。



其中w（t）是一个权重函数。如果w是一个非零连续函数，使得w（x）≥0开（−1,1），那么上述双线性形式确实是正定的。正交族

11.5。线性等轴测（正交变换）

近似理论和物理中使用的多项式是由适当选择权函数w产生的。除了前面的两个例子外，厄米特多项式对应于w（x）=e−x2，拉盖尔多项式对应于w（x）=e−x，雅可比多项式对应于w（x）=（1−x）α（1+x）β，w。ithα，β>−1.正交多项式的综合处理可在Lebedev[111]、Sansone[140]和Andrews、Askey和Roy[3]中找到。

我们也可以证明关于正交空间的下列命题。

提案11.11.给定有限维n≥1的任意非平凡欧几里德空间e，对于维k的任何子空间f，f的正交补f具有维n−k，e=f\_f\_。此外，我们还有f=f。

证据。从命题11.9，子空间f有一些正交基（u1，…，uk）。这个线性独立族（u1，…，uk）可以推广到一个基（u1，…，uk，vk+1，…，vn），通过命题11.10，它可以转化为一个正态基（u1，…，un），其中包含（u1，…，uk）作为f的正态基，现在任何向量w=w1u1+·············+wnun∈e与f正交。i f f w·ui=0，对于每个i，其中1≤i≤k，iff wi=0，对于每个i，其中1≤i≤k。很明显，这表明（uk+1，…，un）是f的基础，因此e=f f，f具有维数n−k。同样，任何向量w=w1u1+··············+wun∈e与f iff w·ui正交。=0，对于每个i，其中k+1≤i≤n，i f f wi=0，对于每个i，其中k+1≤i≤n。因此，（u1，…，uk）是f和f=f的基础。

## 11.5线性等距图（正交变换）

在本节中，我们考虑保留欧几里得范数的欧几里得空间之间的线性映射。这些变换，有时称为刚性运动，在几何学中起着重要作用。

定义11.5.在任意两个非平凡欧几里德空间e和f相同的有限维n下，函数f:e→f是一个正交变换，或线性等距变换，如果它是线性的并且kf（u）k=kuk，对于所有u∈e。

评论：

1. 线性等值线通常被定义为一个线性图，这样

k f（v）−f（u）k=kv−uk，

对于所有u，v∈e，由于映射f是线性的，这两个定义是等价的。第二个定义只关注于保持向量之间的距离。

1. 有时，满足定义11.5的条件的线性映射称为度量映射，线性等值线定义为双射度量映射。

An isometry (without the word linear) is sometimes defined as a function *f* : *E* → *F* (not necessarily linear) such that

k*f*(*v*) − *f*(*u*)k = k*v* − *u*k*,*

for all *u,v* ∈ *E*, i.e., as a function that preserves the distance. This requirement turns out to be very strong. Indeed, the next proposition shows that all these definitions are equivalent when *E* and *F* are of finite dimension, and for functions such that *f*(0) = 0.

**Proposition 11.12.** *Given any two nontrivial Euclidean spaces E and F of the same finite dimension n, for every function f* : *E* → *F, the following properties are equivalent:*

*(1) f is a linear map and* k*f*(*u*)k = k*u*k*, for all u* ∈ *E; (2)* k*f*(*v*) − *f*(*u*)k = k*v* − *u*k*, for all u,v* ∈ *E, and f*(0) = 0*;*

*(3) f*(*u*) · *f*(*v*) = *u* · *v, for all u,v* ∈ *E.*

*Furthermore, such a map is bijective.*

*Proof.* Clearly, (1) implies (2), since in (1) it is assumed that *f* is linear.

Assume that (2) holds. In fact, we shall prove a slightly stronger result. We prove that if

k*f*(*v*) − *f*(*u*)k = k*v* − *u*k

for all *u,v* ∈ *E*, then for any vector *τ* ∈ *E*, the function *g*: *E* → *F* defined such that

*g*(*u*) = *f*(*τ* + *u*) − *f*(*τ*)

for all *u* ∈ *E* is a linear map such that *g*(0) = 0 and (3) holds. Clearly, *g*(0) = *f*(*τ*)−*f*(*τ*) = 0. Note that from the hypothesis

k*f*(*v*) − *f*(*u*)k = k*v* − *u*k

for all *u,v* ∈ *E*, we conclude that

k*g*(*v*) − *g*(*u*)k = k*f*(*τ* + *v*) − *f*(*τ*) − (*f*(*τ* + *u*) − *f*(*τ*))k*,*

= k*f*(*τ* + *v*) − *f*(*τ* + *u*)k*,*

= k*τ* + *v* − (*τ* + *u*)k*,* = k*v* − *u*k*,*

for all *u,v* ∈ *E*. Since *g*(0) = 0, by setting *u* = 0 in

k*g*(*v*) − *g*(*u*)k = k*v* − *u*k*,*