w  
W

Figure 17.4: The basis functions wi0 and wj1  
图17.4：基础函数wi0和wj1

# 17.2 A Two-Dimensional Problem: An Elastic Membrane 17.2二维问题：弹性膜

Consider an elastic membrane attached to a round contour whose projection on the (x1,x2)plane is the boundary Γ of an open, connected, bounded region Ω in the (x1,x2)-plane, as illustrated in Figure 17.5. In other words, we view the membrane as a surface consisting of the set of points (x,z) given by an equation of the form  
假设一个弹性膜附着在一个圆形轮廓上，其在（x1，x2）平面上的投影是（x1，x2）平面中开放、连接、有界区域Ω的边界\_，如图17.5所示。换句话说，我们把膜看作是一个表面，由一个形式方程给出的一组点（x，z）组成。

z = u(x),  
Z=U（X）、

with x = (x1,x2) ∈ Ω, where u: Ω → R is some sufficiently regular function, and we think of u(x) as the vertical displacement of this membrane.  
对于x=（x1，x2）∈Ω，其中u:Ω→r是一些足够规则的函数，我们认为u（x）是膜的垂直位移。

We assume that this membrane is under the action of a vertical force τf(x)dx per surface element in the horizontal plane (where τ is the tension of the membrane). The problem is  
我们假设该膜在水平面上每个表面单元的垂直力τf（x）dx的作用下（其中τ是膜的张力）。问题是

## w17.2. A TWO-DIMENSIONAL PROBLEM: AN ELASTIC MEMBRANE 图17.2。二维问题：弹性膜

*x*

1

*x*

2

Γ

*y*

*g*

(

*y*

)

Ω

*u*

(

*x*

)

*x*

*τ*

*f*

(

*x*

)

*dx*

*dx*

Figure 17.5: An elastic membrane  
图17.5：弹性膜

to find the vertical displacement u as a function of x, for x ∈ Ω. It can be shown (under some assumptions on Ω, Γ, and f), that u(x) is given by a PDE with boundary condition, of the form  
求垂直位移u与x的函数关系，对于x∈Ω。可以证明（在一些关于Ω、\_和f的假设下），u（x）由具有边界条件的PDE给出，形式为

|  |  |
| --- | --- |
| −∆u(x) = f(x), 网络错误 | x ∈ Ω 网络错误 |
| u(x) = g(x), 网络错误 | x ∈ Γ, 网络错误 |

where g: Γ → R represents the height of the contour of the membrane. We are looking for  
其中g：\_→r代表膜轮廓的高度。我们在找

a function u in C2(Ω) ∩ C1(Ω). The operator ∆ is the Laplacian, and it is given by  
C2（Ω）C1（Ω）中的功能U。算符∆是拉普拉斯函数，它由

.  
.

This is an example of a boundary problem, since the solution u of the PDE must satisfy the condition u(x) = g(x) on the boundary of the domain Ω. The above equation is known as Poisson’s equation, and when f = 0 as Laplace’s equation.  
这是一个边界问题的例子，因为PDE的解u必须满足域Ω边界上的条件u（x）=g（x）。上述方程称为泊松方程，当f=0时称为拉普拉斯方程。

It can be proved that if the data f,g and Γ are sufficiently smooth, then the problem has a unique solution.  
可以证明，如果数据f、g和\_足够平滑，那么问题有一个独特的解决方案。

To get a weak formulation of the problem, first we have to make the boundary condition homogeneous, which means that g(x) = 0 on Γ. It turns out that g can be extended to the  
为了得到问题的弱表达式，首先要使边界条件均匀，即g（x）=0 on\_。结果证明G可以扩展到

whole of Ω as some sufficiently smooth function bh, so we can look for a solution of the form u − bh, but for simplicity, let us assume that the contour of Ω lies in a plane parallel to the  
整个Ω是一些足够平滑的函数bh，因此我们可以寻找形式u-bh的解，但为了简单起见，让我们假设Ω的轮廓位于一个平行于

(x1,x2)- plane, so that g = 0. We let V be the subspace of C2(Ω) ∩ C1(Ω) consisting of functions v such that v = 0 on Γ.  
（x1，x2）-平面，因此g=0。我们将v设为c2（Ω）c1（Ω）的子空间，由函数v组成，这样v=0 on\_。

As before, we multiply the PDE by a test function v ∈ V , getting  
如前所述，我们将pde乘以一个测试函数v∈v，得到

−∆u(x)v(x) = f(x)v(x),  
−∆U（x）V（x）=F（x）V（x）、

and we “integrate by parts.” In this case, this means that we use a version of Stokes formula known as Green’s first identity, which says that  
在这种情况下，这意味着我们使用了斯托克斯公式的一个版本，称为格林的第一个恒等式，也就是说

Z Z Z  
Z z

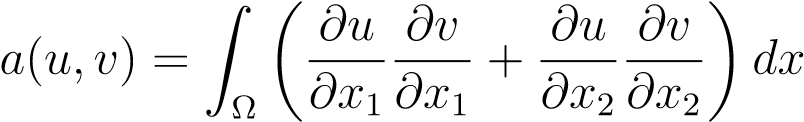
−∆uv dx = (gradu) · (gradv)dx − (gradu) · nvdσ  
−∆uv dx=（梯度u）·（梯度v）dx−（梯度u）·nvdσ

Ω Ω Γ  
欧姆

(where n denotes the outward pointing unit normal to the surface). Because v = 0 on Γ, the integral RΓ drops out, and we get an equation of the form  
（其中n表示垂直于表面的向外指向单位）。因为在\_上v=0，所以积分r\_就消失了，我们得到了形式的方程。

a(u,v) = fe(v) for all v ∈ V,  
a（u，v）=fe（v），对于所有v∈v，

where a is the bilinear form given by  
其中a是双线性形式



and feis the linear form given by Z  
并feis z给出的线性形式

fe(v) = fvdx.  
铁（V）=fvdx。

Ω  
γ

We get the same equation as in section 17.2, but over a set of functions defined on a two-dimensional domain. As before, we can choose a finite-dimensional subspace Va of V and consider the discrete problem with respect to Va. Again, if we pick a basis (w1,...,wn) of Va, a vector u = u1w1 + ··· + unwn is a solution of the Weak Formulation of our problem iff u = (u1,...,un) is a solution of the linear system  
我们得到与第17.2节中相同的方程，但在二维域上定义的一组函数上。和以前一样，我们可以选择v的有限维子空间v a，并考虑与va有关的离散问题。同样，如果我们选择va的基（w1，…，wn），那么向量u=u1w1+········+unwn是我们问题弱公式的解，如果u=（u1，…，un）是线性系统的解。茎

Au = b,  
au=b，

with A = (a(wi,wj)) and b = (fe(wj)). However, the integrals that give the entries in A and b are much more complicated.  
其中a=（a（wi，wj））和b=（fe（wj））。然而，给a和b中的项的积分要复杂得多。

An approach to deal with this problem is the method of finite elements. The idea is to also discretize the boundary curve Γ. If we assume that Γ is a polygonal line, then we can triangulate the domain Ω, and then we consider spaces of functions which are piecewise defined on the triangles of the triangulation of Ω. The simplest functions are piecewise affine and look like tents erected above groups of triangles. Again, we can define base functions with small support, so that the matrix A is tridiagonal by blocks.  
处理这个问题的方法是有限元法。其思想是对边界曲线\_进行离散化。如果假设\_是一条多边形线，那么我们可以对域Ω进行三角化，然后我们考虑在Ω三角化的三角形上分段定义的函数空间。最简单的函数是分段仿射函数，它看起来就像一组三角形上面的帐篷。同样，我们可以用小的支持来定义基函数，这样矩阵A就成了三对角的块。

The finite element method is a vast subject and it is presented in many books of various degrees of difficulty and obscurity. Let us simply state three important requirements of the finite element method:  
有限元法是一门庞大的学科，在许多不同难度和晦涩程度的书籍中都有介绍。让我们简单地说明有限元法的三个重要要求：

1. “Good” triangulations must be found. This in itself is a vast research topic. Delaunay triangulations are good candidates.  
   必须找到“良好”的三角测量。这本身就是一个庞大的研究课题。Delaunay三角测量是很好的候选者。
2. “Good” spaces of functions must be found; typically piecewise polynomials and splines.  
   必须找到“好”的函数空间；通常是分段多项式和样条曲线。
3. “Good” bases consisting of functions will small support must be found, so that integrals can be easily computed and sparse banded matrices arise.  
   由函数组成的“好”基的支持度很小，因此积分很容易计算，并产生稀疏带状矩阵。

We now consider boundary problems where the solution varies with time.  
我们现在考虑边界问题，其中解随时间变化。

# 17.3 Time-Dependent Boundary Problems: The Wave Equation 17.3时变边界问题：波动方程

Consider a homogeneous string (or rope) of constant cross-section, of length L, and stretched (in a vertical plane) between its two ends which are assumed to be fixed and located along the x-axis at x = 0 and at x = L.  
考虑等截面、长度为l的均质绳索（或绳索），并在其两端之间拉伸（垂直平面），假定其固定并沿x轴位于x=0和x=l处。

Figure 17.6: A vibrating string  
图17.6：振动弦

The string is subjected to a transverse force τf(x)dx per element of length dx (where τ is the tension of the string). We would like to investigate the small displacements of the string in the vertical plane, that is, how it vibrates.  
每根长度为dx（其中，τ是管柱的张力）的元件承受横向力τf（x）dx。我们想研究弦在垂直面上的小位移，也就是它是如何振动的。

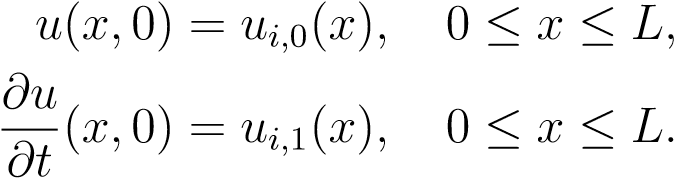
Thus, we seek a function u(x,t) defined for t ≥ 0 and x ∈ [0,L], such that u(x,t) represents the vertical deformation of the string at the abscissa x and at time t.  
因此，我们寻求一个函数u（x，t），定义为t≥0和x∈[0，l]，这样u（x，t）代表弦在横坐标x和时间t处的垂直变形。

It can be shown that u must satisfy the following PDE  
可以看出，u必须满足以下pde

,  
，

with c = pτ/ρ, where ρ is the linear density of the string, known as the one-dimensional wave equation.  
c=pτ/ρ，其中，ρ是弦的线密度，称为一维波动方程。

Furthermore, the initial shape of the string is known at t = 0, as well as the distribution of the initial velocities along the string; in other words, there are two functions ui,0 and ui,1 such that  
此外，弦的初始形状已知为t=0，以及沿弦的初始速度分布；换句话说，有两个函数ui，0和ui，1，这样



For example, if the string is simply released from its given starting position, we have ui,1 = 0. Lastly, because the ends of the string are fixed, we must have  
例如，如果字符串只是从给定的起始位置释放，那么我们有ui，1=0。最后，因为字符串的两端是固定的，所以我们必须

u(0,t) = u(L,t) = 0, t ≥ 0.  
u（0，t）=u（l，t）=0，t≥0。

Consequently, we look for a function u: R+ × [0,L] → R satisfying the following conditions:  
因此，我们寻找一个满足以下条件的函数u:r+×0，l]→r：

,  
，

u(0,t) = u(L,t) = 0, t ≥ 0 (boundary condition), u(x,0) = ui,0(x), 0 ≤ x ≤ L (intitial condition),  
u（0，t）=u（l，t）=0，t≥0（边界条件），u（x，0）=ui，0（x），0≤x≤l（初始条件），

(intitial condition).  
（初始条件）。

This is an example of a time-dependent boundary-value problem, with two initial conditions.  
这是一个时间相关边值问题的例子，有两个初始条件。

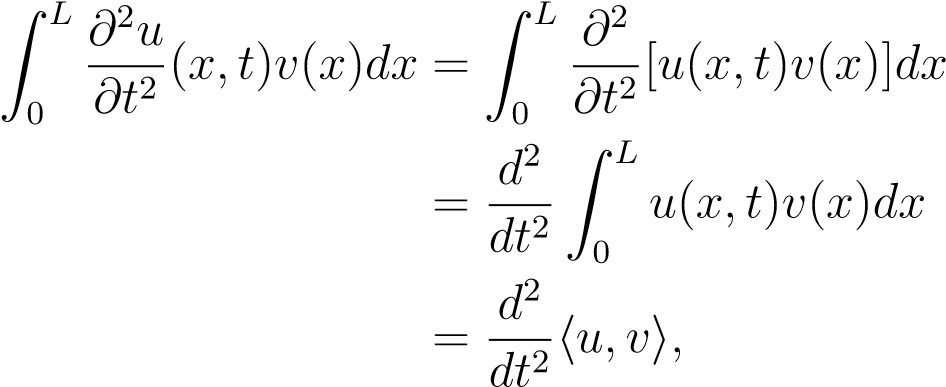
To simplify the problem, assume that f = 0, which amounts to neglecting the effect of gravity. In this case, our PDE becomes  
为了简化这个问题，假设f=0，这等于忽略了重力的影响。在这种情况下，我们的PDE变成

,  
，

Let us try our trick of multiplying by a test function v depending only on x, C1 on [0,L], and such that v(0) = v(L) = 0, and integrate by parts. We get the equation  
让我们试着用一个只依赖于x，c1的测试函数v乘以[0，l]，这样v（0）=v（l）=0，并按部分积分。我们得到了方程

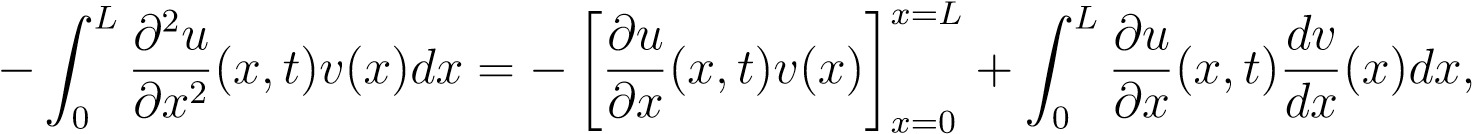
.  
.

For the first term, we get  
第一学期，我们

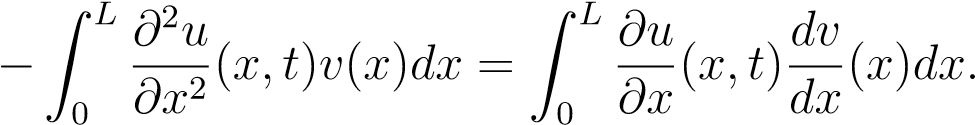


where hu,vi is the inner product in L2([0,L]). The fact that it is legitimate to move ∂2/∂t2 outside of the integral needs to be justified rigorously, but we won’t do it here.  
其中，hu，vi是l2（[0，l]）中的内积。把2/t2移出积分是合法的，这一事实需要严格证明，但我们不会在这里这样做。

For the second term, we get  
第二学期，我们

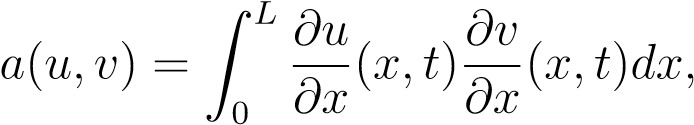


and because v ∈ V , we have v(0) = v(L) = 0, so we obtain  
因为v∈v，我们有v（0）=v（l）=0，所以我们得到



Our integrated equation becomes  
我们的积分方程变成

, for all v ∈ V and all t ≥ 0. It is natural to introduce the bilinear form a: V × V → R given by  
，对于所有v∈v和所有t≥0。引入双线性形式a:v×v→r是自然的，由



where, for every t ∈ R+, the functions u(x,t) and (v,t) belong to V . Actually, we have to replace V by the subspace of the Sobolev space) consisting of the functions such that v(0) = v(L) = 0. Then, the weak formulation (variational formulation) of our problem is this:  
其中，对于每一个t∈r+，函数u（x，t）和（v，t）都属于v。实际上，我们必须用Sobolev空间的子空间替换v），它由函数组成，这样v（0）=v（l）=0。那么，我们问题的弱公式（变分公式）是：

Find a function u ∈ V such that  
求一个函数u∈v，这样

and all t ≥ 0  
且所有t≥0

u(x,0) = ui,0(x), 0 ≤ x ≤ L (intitial condition),  
u（x，0）=ui，0（x），0≤x≤l（初始条件）

(intitial condition).  
（初始条件）。

It can be shown that there is a positive constant α > 0 such that  
可以证明有一个正常数α>0，这样

for all v ∈ V  
对于所有v∈v

(Poincar´e’s inequality), which shows that a is positive definite on V . The above method is known as the method of Rayleigh-Ritz.  
（Poincar'e不等式），表示a在v上是正定的。上述方法称为瑞利-里兹法。

A study of the above equation requires some sophisticated tools of analysis which go far beyond the scope of these notes. Let us just say that there is a countable sequence of solutions with separated variables of the form  
对上述方程的研究需要一些复杂的分析工具，这些工具远远超出了这些注释的范围。我们只要说，有一个可数的解序列，它的形式变量是分开的。

,  
，

called modes (or normal modes). Complete solutions of the problem are series obtained by combining the normal modes, and they are of the form  
被称为模式（或正常模式）。该问题的完全解是由正规模式组合而成的级数，其形式为

,  
，

where the coefficients Ak,Bk are determined from the Fourier series of ui,0 and ui,1.  
其中系数ak，bk由ui，0和ui，1的傅立叶级数确定。

We now consider discrete approximations of our problem. As before, consider a finite dimensional subspace Va of V and assume that we have approximations ua,0 and ua,1 of ui,0 and ui,1. If we pick a basis (w1,...,wn) of Va, then we can write our unknown function u(x,t) as u(x,t) = u1(t)w1 + ··· + un(t)wn,  
我们现在考虑问题的离散近似。如前所述，考虑V的有限维子空间v a，假设我们有近似值ua，0和ua，ui，0和ui，1。如果我们选取VA的基（w1，…，wn），那么我们可以把未知函数u（x，t）写成u（x，t）=u1（t）w1+······+un（t）wn，

where u1,...,un are functions of t. Then, if we write u = (u1,...,un), the discrete version of our problem is  
其中，u1，…，un是t的函数，那么，如果我们写u=（u1，…，un），我们问题的离散版本是

d2u  
D2U

A + Ku = 0,  
a+ku=0，

where A = (hwi,wji) and K = (a(wi,wj)) are two symmetric matrices, called the mass matrix and the stiffness matrix, respectively. In fact, because a and the inner product h−,−i are positive definite, these matrices are also positive definite.  
其中a=（hwi，wji）和k=（a（wi，wj））是两个对称矩阵，分别称为质量矩阵和刚度矩阵。事实上，因为a和内积h−，−i是正定的，所以这些矩阵也是正定的。

We have made some progress since we now have a system of ODE’s, and we can solve it by analogy with the scalar case. So, we look for solutions of the form Ucosωt (or Usinωt), where U is an n-dimensional vector. We find that we should have  
我们已经取得了一些进展，因为我们现在有了一个ODE系统，并且我们可以通过与标量情况进行类比来解决它。因此，我们寻找形式为ucosωt（或usinωt）的解，其中u是一个n维向量。我们发现我们应该

(K − ω2A)Ucosωt = 0,  
（k−ω2a）ucosωt=0，

which implies that ω must be a solution of the equation  
这意味着ω必须是方程的解。

KU = ω2AU.  
ku=ω2au。

Thus, we have to find some λ such that  
因此，我们必须找到一些λ，以便

KU = λAU,  
ku=λau，

a problem known as a generalized eigenvalue problem, since the ordinary eigenvalue problem for K is  
由于k的一般特征值问题是

KU = λU.  
ku=λu.

Fortunately, because A is SPD, we can reduce this generalized eigenvalue problem to a standard eigenvalue problem. A good way to do so is to use a Cholesky decomposition of A as  
幸运的是，由于a是spd，我们可以将这个广义特征值问题简化为标准特征值问题。这样做的一个好方法是使用a的cholesky分解

A = LL>,  
A=ll>，

where L is a lower triangular matrix (see Theorem 7.10). Because A is SPD, it is invertible, so L is also invertible, and  
其中，L是下三角矩阵（见定理7.10）。因为a是spd，它是可逆的，所以l也是可逆的，并且

KU = λAU = λLL>U  
ku=λau=λll>u

yields  
产量

L−1KU = λL>U,  
L−1ku=λl>u，

which can also be written as  
也可以写为

L−1K(L>)−1L>U = λL>U. Then, if we make the change of variable  
L−1K（L>）−1L>U=λL>U。那么，如果我们改变变量

Y = L>U,  
Y=L>U，

using the fact (L>)−1 = (L−1)>, the above equation is equivalent to  
利用事实（l>）-1=（l−1）>，上述方程等于

L−1K(L−1)>Y = λY,  
L−1K（L−1）>Y=λY，

a standard eigenvalue problem for the matrix Kb = L−1K(L−1)>. Furthermore, we know from Section 7.8 that since K is SPD and L−1 is invertible, the matrix Kb = L−1K(L−1)> is also SPD.  
矩阵的标准特征值问题kb=l−1k（l−1）>。此外，我们从第7.8节了解到，由于k是spd，l−1是可逆的，因此矩阵kb=l−1k（l−1）>也是spd。

Consequently, Kb has positive real eigenvalues ( ) (not necessarily distinct) and it can be diagonalized with respect to an orthonormal basis of eigenvectors, say Y1,...,Yn.  
因此，kb具有正的实特征值（）（不一定是不同的），并且它可以相对于特征向量的正态基对角化，例如y1，…，yn。

Then, since Y = L>U, the vectors  
然后，因为y=l>u，向量

Ui = (L>)−1Yi, i = 1,...,n,  
ui=（l>）-1yi，i=1，…，n，

are linearly independent and are solutions of the generalized eigenvalue problem; that is,  
是线性独立的，是广义特征值问题的解；也就是说，

KUi = ωi2AUi, i = 1,...,n.  
kui=ωi2aui，i=1，…，n.

More is true. Because the vectors Y1,...,Yn are orthonormal, and because Yi = L>Ui, from  
更多是真的。因为向量y1，…，yn是正交的，并且因为yi=l>ui，来自

(Yi)>Yj = δij,  
（yi）>yj=δij，

we get  
我们得到

(Ui)>LL>Uj = δij, 1 ≤ i,j ≤ n,  
（ui）>ll>uj=δi j，1≤i，j≤n，

and since A = LL>, this yields  
由于a=ll>，这就产生了

(Ui)>AUj = δij, 1 ≤ i,j ≤ n.  
（ui）>auj=δi j，1≤i，j≤n。

This suggests defining the functions Ui ∈ Va by  
这建议通过定义函数ui∈va

.  
.

Then, it immediate to check that  
那么，马上检查一下

a(Ui,Uj) = (Ui)>AUj = δij,  
a（ui，uj）=（ui）>auj=δij，

which means that the functions (U1,...,Un) form an orthonormal basis of Va for the inner product a. The functions Ui ∈ Va are called modes (or modal vectors).  
这意味着函数（u1，…，un）为内积a形成了va的正交基，函数ui∈va称为模态（或模态向量）。

As a final step, let us look again for a solution of our discrete weak formulation of the problem, this time expressing the unknown solution u(x,t) over the modal basis (U1,...,Un), say  
作为最后一步，让我们再一次寻找问题的离散弱公式的解，这次用模态基（u1，…，un）表示未知解u（x，t），比如

,  
，

where each uej is a function of t. Because  
其中，每个uej都是t的函数，因为

,  
，

if we write u = (u1,...,un) with, we see that  
如果我们写u=（u1，…，un），我们会看到

u ,  
U

so using the fact that  
所以利用这个事实

KUj = ωj2AUj, j = 1,...,n,  
kuj=ωj2auj，j=1，…，n，

the equation  
方程式

d2u  
D2U

A dt2 + Ku = 0  
a dt2+ku=0

yields  
产量

.  
.

Since A is invertible and since (U1,...,Un) are linearly independent, the vectors (AU1,  
由于a是可逆的，并且（u1，…，un）是线性无关的，因此向量（au1，

...,AUn) are linearly independent, and consequently we get the system of n ODEs’  
…，aun）是线性独立的，因此我们得到了n odes的系统。

(uej)00 + ωj2uej = 0, 1 ≤ j ≤ n.  
（uej）00+ωj2uej=0，1≤j≤n。

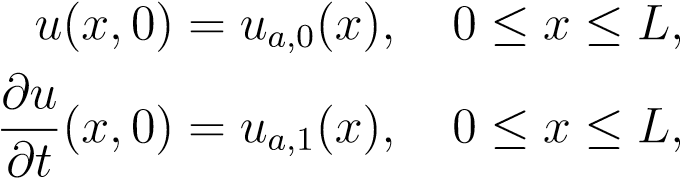
Each of these equation has a well-known solution of the form  
每个方程都有一个众所周知的形式解。

uej = Aj cosωjt + Bj sinωjt.  
uej=aj cosωjt+bj sinωjt。

Therefore, the solution of our approximation problem is given by  
因此，我们的近似问题的解由

,  
，

and the constants Aj,Bj are obtained from the intial conditions  
从初始条件中得到常数aj，bj。



by expressing ua,0 and ua,1 on the modal basis (U1,...,Un). Furthermore, the modal functions (U1,...,Un) form an orthonormal basis of Va for the inner product a.  
通过在模态基础上表达ua，0和ua，1（u1，…，un）。此外，模态函数（u1，…，un）为内积a形成了va的正交基。

If we use the vector space VN0 of piecewise affine functions, we find that the matrices A and K are familar! Indeed,  
如果我们使用分段仿射函数的向量空间vn0，我们会发现矩阵a和k是家族的！的确，

and   
和

To conclude this section, let us discuss briefly the wave equation for an elastic membrane, as described in Section 17.2. This time, we look for a function u: R+ × Ω → R satisfying the following conditions:  
为了总结这一节，让我们简要讨论弹性膜的波动方程，如第17.2节所述。这次，我们寻找一个满足以下条件的函数u:r+×Ω→r：

,  
，

u(x,t) = 0, x ∈ Γ, t ≥ 0 (boundary condition), u(x,0) = ui,0(x), x ∈ Ω (intitial condition), Ω (intitial condition).  
u（x，t）=0，x∈\_，t≥0（边界条件），u（x，0）=ui，0（x），x∈Ω（初始条件），Ω（初始条件）。

Assuming that f = 0, we look for solutions in the subspace V of the Sobolev space consisting of functions v such that v = 0 on Γ. Multiplying by a test function v ∈ V and using Green’s first identity, we get the weak formulation of our problem:  
假设f=0，我们在Sobolev空间的子空间v中寻找由函数v组成的解，这样v=0 on\_。用一个检验函数v∈v乘以格林的第一个恒等式，得到了问题的弱公式：

Find a function u ∈ V such that  
求一个函数u∈v，这样

and all t ≥ 0  
且所有t≥0

u(x,0) = ui,0(x), x ∈ Ω (intitial condition),  
u（x，0）=ui，0（x），x∈Ω（初始条件）

Ω (intitial condition),  
Ω（初始状态）

where a: V × V → R is the bilinear form given by  
其中a:v×v→r为双线性形式，由

and   
和

As usual, we find approximations of our problem by using finite dimensional subspaces Va of V . Picking some basis (w1,...,wn) of Va, and triangulating Ω, as before, we obtain the equation  
和往常一样，我们用有限维子空间v的va来近似我们的问题。选取Va的一些基（w1，…，wn），然后像以前一样进行三角处理，我们得到了方程。

d2u  
D2U

A dt2 + Ku = 0,  
dt2+ku=0，

,  
，

where A = (hwi,wji) and K = (a(wi,wj)) are two symmetric positive definite matrices.  
其中a=（hwi，wji）和k=（a（wi，wj））是两个对称正定矩阵。

In principle, the problem is solved, but, it may be difficult to find good spaces Va, good triangulations of Ω, and good bases of Va, to be able to compute the matrices A and K, and to ensure that they are sparse.  
原则上解决了这个问题，但要计算矩阵a和k并确保它们是稀疏的，可能很难找到好的空间va、Ω的良好三角和va的良好基。

Chapter 18  
第十八章

# Graphs and Graph Laplacians; Basic Facts 图形和拉普拉斯图形.基本事实

In this chapter and the next we present some applications of linear algebra to graph theory. Graphs (undirected and directed) can be defined in terms of various matrices (incidence and adjacency matrices), and various connectivity properties of graphs are captured by properties of these matrices. Another very important matrix is associated with a (undirected) graph: the graph Laplacian. The graph Laplacian is symmetric positive definite, and its eigenvalues capture some of the properties of the underlying graph. This is a key fact that is exploited in graph clustering methods, the most powerful being the method of normalized cuts due to Shi and Malik [155]. This chapter and the next constitute an introduction to algebraic and spectral graph theory. We do not discuss normalized cuts, but we discuss graph drawings. Thorough presentations of algebraic graph theory can be found in Godsil and Royle [77] and Chung [39].  
在本章和下一章中，我们将介绍线性代数在图论中的一些应用。图（无向和有向）可以用各种矩阵（关联矩阵和邻接矩阵）来定义，图的各种连通性由这些矩阵的性质来捕获。另一个非常重要的矩阵与（无向）图有关：拉普拉斯图。拉普拉斯图是对称正定的，它的特征值捕获了底层图的一些性质。这是一个关键的事实，在图聚类方法中得到了充分利用，其中最强大的是由Shi和Malik[155]提出的归一化切割方法。本章和下一章将介绍代数和谱图理论。我们不讨论标准化切割，但讨论图形绘制。代数图论的详细介绍可在Godsil、Royle[77]和Chung[39]中找到。

We begin with a review of basic notions of graph theory. Even though the graph Laplacian is fundamentally associated with an undirected graph, we review the definition of both directed and undirected graphs. For both directed and undirected graphs, we define the degree matrix D, the incidence matrix B, and the adjacency matrix A. Then we define a weighted graph. This is a pair (V,W), where V is a finite set of nodes and W is a m × m symmetric matrix with nonnegative entries and zero diagonal entries (where m = |V |).  
我们首先回顾了图论的基本概念。尽管拉普拉斯图基本上与无向图有关，但我们回顾了有向图和无向图的定义。对于有向图和无向图，我们定义了度矩阵D、关联矩阵B和邻接矩阵A，然后定义了加权图。这是一对（v，w），其中v是一组有限的节点，w是一个m×m对称矩阵，具有非负项和零对角项（其中m=v）。

For every node vi ∈ V , the degree d(vi) (or di) of vi is the sum of the weights of the edges adjacent to vi:  
对于每个节点vi∈v，vi的阶数d（vi）（或di）是与vi相邻的边的权重之和：

.  
.

The degree matrix is the diagonal matrix  
度矩阵是对角矩阵

D = diag(d1,...,dm).  
d=诊断（d1，…，dm）。

The notion of degree is illustrated in Figure 18.1. Then we introduce the (unnormalized)  
度的概念如图18.1所示。然后我们介绍（非规范化的）

579  
五百七十九

Figure 18.1: Degree of a node.  
图18.1：节点的度数。

graph Laplacian L of a directed graph G in an “old-fashion” way, by showing that for any orientation of a graph G,  
有向图G的拉普拉斯L，用“旧的方式”表示，对于图G的任何方向，

BB> = D − A = L  
bb>=d−a=l

is an invariant. We also define the (unnormalized) graph Laplacian L of a weighted graph G = (V,W) as L = D−W. We show that the notion of incidence matrix can be generalized to weighted graphs in a simple way. For any graph Gσ obtained by orienting the underlying graph of a weighted graph G = (V,W), there is an incidence matrix Bσ such that  
是不变量。我们还将加权图g=（v，w）的（非规范化）拉普拉斯L定义为L=d−w。我们证明了关联矩阵的概念可以简单地推广到加权图。对于通过将加权图G=（V，W）的基础图定向而得到的任何图Gσ，存在一个关联矩阵Bσ，使得

Bσ(Bσ)> = D − W = L.  
bσ（bσ）>=d−w=l。

We also prove that  
我们也证明了

for all x ∈ Rm.  
对于所有x∈rm。

Consequently, x>Lx does not depend on the diagonal entries in W, and if wij ≥ 0 for all i,j ∈ {1,...,m}, then L is positive semidefinite. Then if W consists of nonnegative entries, the eigenvalues 0 = λ1 ≤ λ2 ≤ ... ≤ λm of L are real and nonnegative, and there is an orthonormal basis of eigenvectors of L. We show that the number of connected components of the graph G = (V,W) is equal to the dimension of the kernel of L, which is also equal to the dimension of the kernel of the transpose (Bσ)> of any incidence matrix Bσ obtained by orienting the underlying graph of G.  
因此，x>lx不依赖于w中的对角线项，如果所有i，j∈1，…，m的wij≥0，则l为正半定。如果w由非负项组成，则特征值0=λ1≤λ2≤…L的λm是实的和非负的，L的特征向量有一个正交基，我们证明了图G的连通分量的个数（v，w）等于L的核的维数，也等于任意一个内点的转置（bσ）>的核的维数。通过确定G的下垫图的方向得到的Ence矩阵bσ。

We also define the normalized graph Laplacians Lsym and Lrw, given by  
我们还定义了标准化的拉普拉斯图lsym和lrw，由

Lsym = D−1/2LD−1/2 = I − D−1/2WD−1/2  
LSYM=D−1/2LD−1/2=I−D−1/2WD−1/2

Lrw = D−1L = I − D−1W,  
lrw=d−1l=i−d−1w，

and prove some simple properties relating the eigenvalues and the eigenvectors of L, Lsym and Lrw. These normalized graph Laplacians show up when dealing with normalized cuts.  
并证明了L、LSYM和LRW的特征值和特征向量的一些简单性质。这些归一化图形拉普拉斯出现在处理归一化切割。

Next, we turn to graph drawings (Chapter 19). Graph drawing is a very attractive application of so-called spectral techniques, which is a fancy way of saying that that eigenvalues and eigenvectors of the graph Laplacian are used. Furthermore, it turns out that graph clustering using normalized cuts can be cast as a certain type of graph drawing.  
接下来，我们来看看图表（第19章）。图的绘制是所谓的谱技术的一个非常有吸引力的应用，这是一种奇特的说法，即图的拉普拉斯特征值和特征向量被使用。此外，结果表明，使用归一化切割的图形聚类可以被转换为某种类型的图形绘制。

Given an undirected graph G = (V,E), with |V | = m, we would like to draw G in Rn for n (much) smaller than m. The idea is to assign a point ρ(vi) in Rn to the vertex vi ∈ V , for every vi ∈ V , and to draw a line segment between the points ρ(vi) and ρ(vj). Thus, a graph drawing is a function ρ: V → Rn.  
给出一个无向图G=（v，e），用v\_=m，我们想在RN中画G，对于小于m的n（多），其思想是将RN中的点ρ（vi）赋给顶点vi∈v，对于每个vi∈v，并在点ρ（vi）和ρ（vj）之间画一条直线段。因此，图形绘制是函数ρ：v→rn。

We define the matrix of a graph drawing ρ (in Rn) as a m × n matrix R whose ith row consists of the row vector ρ(vi) corresponding to the point representing vi in Rn. Typically, we want n < m; in fact n should be much smaller than m.  
我们将图的矩阵ρ（在r n中）定义为m×n矩阵r，其中第i行由行向量ρ（vi）组成，该行向量与在rn中表示vi的点相对应。通常，我们需要n<m；实际上n应该比m小得多。

Since there are infinitely many graph drawings, it is desirable to have some criterion to decide which graph is better than another. Inspired by a physical model in which the edges are springs, it is natural to consider a representation to be better if it requires the springs to be less extended. We can formalize this by defining the energy of a drawing R by  
由于有无限多的图形绘制，因此需要有某种标准来决定哪个图形优于另一个图形。从一个边缘是弹簧的物理模型中得到启发，如果一个表示法要求弹簧的延伸度较小，那么它自然会考虑更好。我们可以通过定义绘图r的能量

,  
，

where ρ(vi) is the ith row of R and kρ(vi) − ρ(vj)k2 is the square of the Euclidean length of the line segment joining ρ(vi) and ρ(vj).  
其中，ρ（vi）是r的第i行，kρ（vi）−ρ（vj）k2是连接ρ（vi）和ρ（vj）的线段欧几里得长度的平方。

Then “good drawings” are drawings that minimize the energy function E. Of course, the trivial representation corresponding to the zero matrix is optimum, so we need to impose extra constraints to rule out the trivial solution.  
那么“好的图”就是能量函数E最小化的图，当然，对应于零矩阵的平凡表示是最优的，所以我们需要施加额外的约束来排除平凡解。

We can consider the more general situation where the springs are not necessarily identical. This can be modeled by a symmetric weight (or stiffness) matrix W = (wij), with wij ≥ 0. In this case, our energy function becomes  
我们可以考虑弹簧不一定相同的更一般的情况。这可以通过对称重量（或刚度）矩阵w=（wij）来建模，wij≥0。在这种情况下，我们的能量函数变成

.  
.

Following Godsil and Royle [77], we prove that  
根据Godsil和Royle[77]，我们证明

E(R) = tr(R>LR),  
e（r）=tr（r>lr）

where  
哪里

L = D − W,  
L=D-W，

is the familiar unnormalized Laplacian matrix associated with W, and where D is the degree matrix associated with W.  
是熟悉的与w有关的非正规拉普拉斯矩阵，其中d是与w有关的度矩阵。

It can be shown that there is no loss in generality in assuming that the columns of R are pairwise orthogonal and that they have unit length. Such a matrix satisfies the equation R>R = I and the corresponding drawing is called an orthogonal drawing. This condition also rules out trivial drawings.  
结果表明，假设R的列是成对正交的，并且它们具有单位长度，则不存在一般性的损失。这样的矩阵满足方程r>r=i，相应的图称为正交图。这种情况也排除了一些琐碎的绘图。

Then we prove the main theorem about graph drawings (Theorem 19.2), which essentially says that the matrix R of the desired graph drawing is constituted by the n eigenvectors of L associated with the smallest nonzero n eigenvalues of L. We give a number examples of graph drawings, many of which are borrowed or adapted from Spielman [158].  
然后，我们证明了关于图的主要定理（定理19.2），它实质上表示所需图的矩阵R是由L的n个特征向量与L的最小非零n个特征值构成的，我们给出了一些图的例子，其中许多是借用或改编自斯皮尔曼[158]。

## 18.1 Directed Graphs, Undirected Graphs, Incidence Matrices, Adjacency Matrices, Weighted Graphs 18.1有向图、无向图、关联矩阵、邻接矩阵、加权图

Definition 18.1. A directed graph is a pair G = (V,E), where V = {v1,...,vm} is a set of nodes or vertices, and E ⊆ V × V is a set of ordered pairs of distinct nodes (that is, pairs (u,v) ∈ V × V with u =6 v), called edges. Given any edge e = (u,v), we let s(e) = u be the source of e and t(e) = v be the target of e.  
定义18.1.有向图是一对g=（v，e），其中v=v1，…，vm是一组节点或顶点，e v×v是一组有序的不同节点对（即对（u，v）v×v和u=6v），称为边。给定任意边e=（u，v），我们让s（e）=u是e的源，t（e）=v是e的目标。

Remark: Since an edge is a pair (u,v) with u =6 v, self-loops are not allowed. Also, there is at most one edge from a node u to a node v. Such graphs are sometimes called simple graphs.  
注：由于边缘是U=6V的对（U，V），因此不允许出现自循环。此外，从节点U到节点V最多有一个边。这种图有时称为简单图。

An example of a directed graph is shown in Figure 18.2.  
有向图的示例如图18.2所示。

**v**

4

**v**

5

**v**

1

**v**

2

**v**

3

**e**

1

**e**

7

**e**

2

**e**

3

**e**

4

**e**

5

**e**

6

Figure 18.2: Graph G1.  
图18.2：图G1。

Definition 18.2. For every node v ∈ V , the degree d(v) of v is the number of edges leaving or entering v:  
定义18.2.对于每个节点v∈v，v的阶数d（v）是离开或进入v的边数：

d(v) = |{u ∈ V | (v,u) ∈ E or (u,v) ∈ E}|.  
d（v）=u∈v（v，u）∈e或（u，v）∈e。

We abbreviate d(vi) as di. The degree matrix, D(G), is the diagonal matrix  
我们把d（vi）缩写为di。度矩阵d（g）是对角矩阵

D(G) = diag(d1,...,dm).  
d（g）=diag（d1，…，dm）。

For example, for graph G1, we have  
例如，对于图g1，我们有

2 0 0 0 0  
2 0 0 0 0\_

0 4 0 0 0  
0 4 0 0 0\_

D(G1) = 0 0 3 0 0.  
D（g1）=0 0 3 0 0。

   
 

0 0 0 3 0  
0 0 3 0\_

0 0 0 0 2  
0 0 0 0 2

Unless confusion arises, we write D instead of D(G).  
除非出现混淆，我们写D而不是D（G）。

Definition 18.3. Given a directed graph G = (V,E), for any two nodes u,v ∈ V , a path from u to v is a sequence of nodes (v0,v1,...,vk) such that v0 = u, vk = v, and (vi,vi+1) is an edge in E for all i with 0 ≤ i ≤ k − 1. The integer k is the length of the path. A path is closed if u = v. The graph G is strongly connected if for any two distinct nodes u,v ∈ V , there is a path from u to v and there is a path from v to u.  
定义18.3.给定有向图g=（v，e），对于任意两个节点u，v∈v，从u到v的路径是节点序列（v0，v1，…，v k），因此v0=u，vk=v，并且（v i，vi+1）是e中所有i的边，0≤i≤k−1。整数k是路径的长度。当u=v时，路径是封闭的，图G是强连通的，如果对于任意两个不同的节点u，v∈v，有一条从u到v的路径，有一条从v到u的路径。

Remark: The terminology walk is often used instead of path, the word path being reserved to the case where the nodes vi are all distinct, except that v0 = vk when the path is closed.  
备注：术语walk通常被用来代替path，单词path被保留到节点vi都是不同的情况下，除了路径关闭时v0=vk。

The binary relation on V ×V defined so that u and v are related iff there is a path from u to v and there is a path from v to u is an equivalence relation whose equivalence classes are called the strongly connected components of G.  
v×v上的二元关系定义为u和v是相关的，如果存在从u到v的路径，并且存在从v到u的路径，则为等价关系，其等价类称为g的强连通分量。

Definition 18.4. Given a directed graph G = (V,E), with V = {v1,...,vm}, if E = {e1,...,en}, then the incidence matrix B(G) of G is the m×n matrix whose entries bij are given by  
定义18.4.给定有向图g=（v，e），其中v=v1，…，v m，如果e=e1，…，e n，则g的关联矩阵b（g）是m×n矩阵，其条目bij由下式给出：

  
γ

+1 if s(ej) = vi bij = −1 if t(ej) = vi 0 otherwise.  
+1，如果s（ej）=vi bij=-1，如果t（ej）=vi 0，否则。

Here is the incidence matrix of the graph G1:  
图g1的关联矩阵如下：

.  
.

Observe that every column of an incidence matrix contains exactly two nonzero entries, +1 and −1. Again, unless confusion arises, we write B instead of B(G).  
注意，一个关联矩阵的每一列正好包含两个非零项，+1和−1。同样，除非出现混淆，我们写B而不是B（G）。

When a directed graph has m nodes v1,...,vm and n edges e1,...,en, a vector x ∈ Rm can be viewed as a function x: V → R assigning the value xi to the node vi. Under this interpretation, Rm is viewed as RV . Similarly, a vector y ∈ Rn can be viewed as a function  
当有向图有M个节点V1，…，VM和N边E1，…，EN时，向量X \* RM可以被看作是函数X：V～R，将值XI赋值给节点VI。同样，向量y∈rn可以看作是一个函数。

*v*

4

*v*

5

*v*

1

*v*

2

*v*

3

*a*

*g*

*b*

*c*

*d*

*e*

*f*

Figure 18.3: The undirected graph G2.  
图18.3：无向图g2。

in RE. This point of view is often useful. For example, the incidence matrix B can be interpreted as a linear map from RE to RV , the boundary map, and B> can be interpreted as a linear map from RV to RE, the coboundary map.  
在RE中。这种观点通常是有用的。例如，关联矩阵b可以解释为从re到rv的线性映射，边界映射，和b>可以解释为从rv到re的线性映射，共边界映射。

Remark: Some authors adopt the opposite convention of sign in defining the incidence matrix, which means that their incidence matrix is −B.  
注：有些作者在定义关联矩阵时采用了符号的相反约定，这意味着他们的关联矩阵是−b。

Undirected graphs are obtained from directed graphs by forgetting the orientation of the edges.  
无向图是通过忽略边的方向从有向图得到的。

Definition 18.5. A graph (or undirected graph) is a pair G = (V,E), where V = {v1,...,vm} is a set of nodes or vertices, and E is a set of two-element subsets of V (that is, subsets {u,v}, with u,v ∈ V and u =6 v), called edges.  
定义18.5.图（或无向图）是一对g=（v，e），其中v=v1，…，vm是一组节点或顶点，e是v的两个元素子集的集合（即子集u，v，具有u，v∈v和u=6v），称为边。

Remark: Since an edge is a set {u,v}, we have u =6 v, so self-loops are not allowed. Also, for every set of nodes {u,v}, there is at most one edge between u and v. As in the case of directed graphs, such graphs are sometimes called simple graphs.  
注：由于边是一组u，v，我们有u=6v，所以不允许自循环。而且，对于每一组节点u，v，u和v之间最多有一个边。在有向图的情况下，这种图有时称为简单图。

An example of a graph is shown in Figure 18.3.  
图18.3显示了一个图表的示例。

Definition 18.6. For every node v ∈ V , the degree d(v) of v is the number of edges incident to v:  
定义18.6.对于每个节点v∈v，v的阶数d（v）是v的边数：

d(v) = |{u ∈ V | {u,v} ∈ E}|.  
d（v）=u∈v u，v∈e。

The degree matrix D(G) (or simply, D) is defined as in Definition 18.2.  
度矩阵d（g）（或简单地说，d）的定义见定义18.2。

Definition 18.7. Given a (undirected) graph G = (V,E), for any two nodes u,v ∈ V , a path from u to v is a sequence of nodes (v0,v1,...,vk) such that v0 = u, vk = v, and {vi,vi+1} is an edge in E for all i with 0 ≤ i ≤ k − 1. The integer k is the length of the path. A path is closed if u = v. The graph G is connected if for any two distinct nodes u,v ∈ V , there is a path from u to v.  
定义18.7.给定（无向）图g=（v，e），对于任意两个节点u，v∈v，从u到v的路径是节点序列（v0，v1，…，v k），使得v0=u，vk=v，和v i，vi+1是e中0≤i≤k−1的所有i的边。整数k是路径的长度。当u=v时，路径是封闭的。图g是连通的，如果对于任意两个不同的节点u，v∈v，有一条从u到v的路径。

Remark: The terminology walk or chain is often used instead of path, the word path being reserved to the case where the nodes vi are all distinct, except that v0 = vk when the path is closed.  
备注：术语walk或chain通常被用来代替path，单词path被保留到节点vi都不同的情况下，除了v0=vk（路径关闭时）。

The binary relation on V ×V defined so that u and v are related iff there is a path from u to v is an equivalence relation whose equivalence classes are called the connected components of G.  
定义了V×V上的二元关系，使U和V是相关的，如果有一条从U到V的路径是一个等价关系，其等价类称为G的连通分量。

The notion of incidence matrix for an undirected graph is not as useful as in the case of directed graphs  
无向图的关联矩阵的概念不如有向图的概念有用。

Definition 18.8. Given a graph G = (V,E), with V = {v1,...,vm}, if E = {e1,...,en}, then the incidence matrix B(G) of G is the m × n matrix whose entries bij are given by  
定义18.8.给定一个图g=（v，e），其中v=v1，…，v m，如果e=e1，…，e n，则g的关联矩阵b（g）是m×n矩阵，其条目bij由下式给出：

(  
（

bij = +1 if ej = {vi,vk} for some k 0 otherwise.  
如果ej=vi，则bij=+1，对于某些k 0，则为vk。

Unlike the case of directed graphs, the entries in the incidence matrix of a graph (undirected) are nonnegative. We usually write B instead of B(G).  
与有向图的情况不同，图（无向）的关联矩阵中的条目是非负的。我们通常写B而不是B（G）。

Definition 18.9. If G = (V,E) is a directed or an undirected graph, given a node u ∈ V , any node v ∈ V such that there is an edge (u,v) in the directed case or {u,v} in the undirected case is called adjacent to u, and we often use the notation  
定义18.9.如果g=（v，e）是有向图或无向图，给定一个节点u∈v，任意节点v∈v，使得有向情况下有一个边（u，v），或在无向情况下有一个边（u，v），我们通常使用符号“邻近u”来表示。

u ∼ v.  
U～V。

Observe that the binary relation ∼ is symmetric when G is an undirected graph, but in general it is not symmetric when G is a directed graph.  
注意，当g是无向图时，二进制关系～是对称的，但一般来说，当g是有向图时，它不是对称的。

The notion of adjacency matrix is basically the same for directed or undirected graphs.  
有向图和无向图的邻接矩阵的概念基本相同。

Definition 18.10. Given a directed or undirected graph G = (V,E), with V = {v1,...,vm}, the adjacency matrix A(G) of G is the symmetric m × m matrix (aij) such that  
定义18.10.给定有向或无向图g=（v，e），对于v=v1，…，v m，g的邻接矩阵a（g）是对称m×m矩阵（aij），这样

1. If G is directed, then  
   如果G被指示，那么

(  
（

aij = 1 if there is some edge (vi,vj) ∈ E or some edge (vj,vi) ∈ E 0 otherwise.  
如果有边（vi，vj）∈e或某边（vj，vi）∈e 0，则aij=1。

1. Else if G is undirected, then  
   否则，如果g是无向的，那么

(  
（

aij = 1 if there is some edge {vi,vj} ∈ E 0 otherwise.  
如果有边vi，则aij=1，否则vj∈e 0。

As usual, unless confusion arises, we write A instead of A(G). Here is the adjacency matrix of both graphs G1 and G2:  
像往常一样，除非出现混淆，否则我们写A而不是A（G）。这是图g1和g2的邻接矩阵：

0 1 1 0 0  
0 1 1 0\_

1 0 1 1 1  
1 0 1 1 1\_

A = 1 1 0 1 0.  
A=1 1 0 1 0。

   
 

0 1 1 0 1  
0 1 1 0 1\_

0 1 0 1 0  
0 1 0 1 0

If G = (V,E) is an undirected graph, the adjacency matrix A of G can be viewed as a linear map from RV to RV , such that for all x ∈ Rm, we have  
如果g=（v，e）是一个无向图，g的邻接矩阵a可以看作是从RV到RV的线性映射，这样对于所有x∈rm，我们得到

(Ax)i = Xxj;  
（ax）i=xxj；

j∼i  
J i

that is, the value of Ax at vi is the sum of the values of x at the nodes vj adjacent to vi. The adjacency matrix can be viewed as a diffusion operator. This observation yields a geometric interpretation of what it means for a vector x ∈ Rm to be an eigenvector of A associated with some eigenvalue λ; we must have  
也就是说，x在vi处的值是x在与vi相邻的节点vj处的值之和。邻接矩阵可视为扩散算子。这个观察给出了向量x∈rm是与某个特征值λ相关的a的特征向量的几何解释；我们必须

λxi = Xxj, i = 1,...,m,  
Xxj＝1，…，m，

j∼i  
J i

which means that the the sum of the values of x assigned to the nodes vj adjacent to vi is equal to λ times the value of x at vi.  
这意味着分配给邻近vi节点vj的x值之和等于λ乘以vi处x值。

Definition 18.11. Given any undirected graph G = (V,E), an orientation of G is a function σ: E → V × V assigning a source and a target to every edge in E, which means that for every edge {u,v} ∈ E, either σ({u,v}) = (u,v) or σ({u,v}) = (v,u). The oriented graph Gσ obtained from G by applying the orientation σ is the directed graph Gσ = (V,Eσ), with Eσ = σ(E).  
定义18.11.给定任意无向图g=（v，e），g的方向是一个函数σ：e→v×v，将一个源和一个目标赋给e中的每一条边，这意味着对于每一条边u，v∈e，要么是σ（u，v）=（u，v）或σ（u，v）=（v，u）。应用方向σ从g得到的定向图gσ是有向图gσ=（v，eσ），其中eσ=σ（e）。

The following result shows how the number of connected components of an undirected graph is related to the rank of the incidence matrix of any oriented graph obtained from G.  
下面的结果显示了无向图的连通分量的数目与从G得到的任何有向图的关联矩阵的秩是如何相关的。

Proposition 18.1. Let G = (V,E) be any undirected graph with m vertices, n edges, and c connected components. For any orientation σ of G, if B is the incidence matrix of the oriented graph Gσ, then c = dim(Ker(B>)), and B has rank m − c. Furthermore, the nullspace of B> has a basis consisting of indicator vectors of the connected components of G; that is, vectors (z1,...,zm) such that zj = 1 iff vj is in the ith component Ki of G, and zj = 0 otherwise.  
提案18.1.设g=（v，e）为任意具有m个顶点、n个边和c个连通分量的无向图。对于g的任何方向σ，如果b是有向图gσ的关联矩阵，则c=dim（ker（b>），b的秩为m−c。此外，b>的零空间有一个基，由g的连接分量的指示向量组成；也就是说，向量（z1，…，zm），使得zj=1 iff vj在g的第i个分量ki中，否则zj=0。

Proof. (After Godsil and Royle [77], Section 8.3). The fact that rank(B) = m − c will be proved last.  
证据。（在Godsil和Royle[77]之后，第8.3节）。排名（b）=m-c的事实将最后证明。

Let us prove that the kernel of B> has dimension c. A vector z ∈ Rm belongs to the kernel of B> iff B>z = 0 iff z>B = 0. In view of the definition of B, for every edge {vi,vj} of G, the column of B corresponding to the oriented edge σ({vi,vj}) has zero entries except for a +1 and a −1 in position i and position j or vice-versa, so we have  
让我们证明b>的核具有维数c，向量z∈rm属于b>iff b>z=0 iff z>b=0的核。根据b的定义，对于g的每一个边vi，vj，b的列对应于定向边σ（vi，vj），除了位置i和位置j中的a+1和a−1外，其余都是零项，因此我们有

zi = zj.  
zi=zj.

An easy induction on the length of the path shows that if there is a path from vi to vj in G (unoriented), then zi = zj. Therefore, z has a constant value on any connected component of  
对路径长度的一个简单归纳表明，如果在g中有一条从vi到vj的路径（无方向），那么zi=zj。因此，z在

G. It follows that every vector z ∈ Ker(B>) can be written uniquely as a linear combination  
g.由此可知，每个向量z∈ker（b>）都可以唯一地写成一个线性组合。

z = λ1z1 + ··· + λczc,  
Z=λ1z1+····+λczc，

where the vector zi corresponds to the ith connected component Ki of G and is defined such that  
其中，矢量zi对应于g的第i个连通分量ki，并且定义为

(  
（

zji = 1 iff vj ∈ Ki  
zji=1 iff vj∈ki

0 otherwise.  
否则为0。

This shows that dim(Ker(B>)) = c, and that Ker(B>) has a basis consisting of indicator vectors.  
这表明dim（ker（b>）=c，而ker（b>）有一个由指示器向量组成的基。

Since B> is a n × m matrix, we have  
因为b>是一个n×m矩阵，我们有

m = dim(Ker(B>)) + rank(B>),  
m=dim（ker（b>）+等级（b>），

and since we just proved that dim(Ker(B>)) = c, we obtain rank(B>) = m − c. Since B and B> have the same rank, rank(B) = m − c, as claimed.   
由于我们刚刚证明了dim（ker（b>）=c，我们得到了秩（b>）=m−c，因为b和b>具有相同的秩，如所述，秩（b）=m−c。

Definition 18.12. Following common practice, we denote by 1 the (column) vector (of dimension m) whose components are all equal to 1.  
定义18.12.按照惯例，我们用1表示（列）向量（维数m），其分量都等于1。

Since every column of B contains a single +1 and a single −1, the rows of B> sum to zero, which can be expressed as  
因为B的每一列都包含一个+1和一个-1，所以B>的行和为零，可以表示为

B>1 = 0.  
b>1=0。

According to Proposition 18.1, the graph G is connected iff B has rank m−1 iff the nullspace of B> is the one-dimensional space spanned by 1.  
根据命题18.1，图G是连通的，如果b的秩为m-1，如果b>的空空间是1的一维空间。

In many applications, the notion of graph needs to be generalized to capture the intuitive idea that two nodes u and v are linked with a degree of certainty (or strength). Thus, we assign a nonnegative weight wij to an edge {vi,vj}; the smaller wij is, the weaker is the link (or similarity) between vi and vj, and the greater wij is, the stronger is the link (or similarity) between vi and vj.  
在许多应用中，需要对图的概念进行广义化，以获取直观的观点，即两个节点u和v与一定程度的确定性（或强度）相关联。因此，我们将非负权重wij赋给边vi，vj；wij越小，vi和vj之间的链接（或相似性）越弱，wij越大，vi和vj之间的链接（或相似性）越强。

Definition 18.13. A weighted graph is a pair G = (V,W), where V = {v1,...,vm} is a set of nodes or vertices, and W is a symmetric matrix called the weight matrix, such that wij ≥ 0 for all i,j ∈ {1,...,m}, and wii = 0 for i = 1,...,m. We say that a set {vi,vj} is an edge iff wij > 0. The corresponding (undirected) graph (V,E) with E = {{vi,vj} | wij > 0}, is called the underlying graph of G.  
定义18.13.加权图是一对g=（v，w），其中v=v1，…，v m是一组节点或顶点，w是称为权矩阵的对称矩阵，这样w i j对于所有i，j 1，…，m，wii=0对于i=1，…，m都等于0。我们说，集vi，vj是边iff wij>0。对应的（无向）图（v，e），其中e=vi，vj wij>0，称为G的底层图。

Remark: Since wii = 0, these graphs have no self-loops. We can think of the matrix W as a generalized adjacency matrix. The case where wij ∈ {0,1} is equivalent to the notion of a graph as in Definition 18.5.  
注：由于wii=0，这些图没有自循环。我们可以把矩阵w看作一个广义邻接矩阵。wij 0,1等于定义18.5中的图的概念的情况。

We can think of the weight wij of an edge {vi,vj} as a degree of similarity (or affinity) in an image, or a cost in a network. An example of a weighted graph is shown in Figure 18.4. The thickness of an edge corresponds to the magnitude of its weight.  
我们可以将边vi、vj的权重wij视为图像中的相似度（或相似性），或网络中的成本。加权图的示例如图18.4所示。边缘的厚度与它的重量大小相对应。

Figure 18.4: A weighted graph.  
图18.4：加权图。

Definition 18.14. Given a weighted graph G = (V,W), for every node vi ∈ V , the degree d(vi) of vi is the sum of the weights of the edges adjacent to vi:  
定义18.14.给定加权图g=（v，w），对于每个节点vi∈v，vi的阶数d（vi）是与vi相邻的边的权重之和：

.  
.

Note that in the above sum, only nodes vj such that there is an edge {vi,vj} have a nonzero contribution. Such nodes are said to be adjacent to vi, and we write vi ∼ vj. The degree matrix D(G) (or simply, D) is defined as before, namely by D(G) = diag(d(v1),...,d(vm)).  
请注意，在上述总和中，只有节点vj具有非零贡献，这样就有一个边vi，vj。这样的节点被称为与vi相邻，我们写vi～vj。度矩阵d（g）（或简单地说，d）定义为之前，即d（g）=diag（d（v1），…，d（vm））。

The weight matrix W can be viewed as a linear map from RV to itself. For all x ∈ Rm, we have  
重量矩阵w可以看作是从RV到自身的线性映射。对于所有x∈rm，我们有

(Wx)i = Xwijxj;  
（wx）i=xwijxj；

j∼i  
J i

that is, the value of Wx at vi is the weighted sum of the values of x at the nodes vj adjacent to vi.  
也就是说，wx在vi处的值是x在邻近vi的节点vj处的值的加权和。

Observe that W1 is the (column) vector (d(v1),...,d(vm)) consisting of the degrees of the nodes of the graph.  
观察w1是（列）向量（d（v1），…，d（vm）），由图中节点的度数组成。

We now define the most important concept of this chapter: the Laplacian matrix of a graph. Actually, as we will see, it comes in several flavors.  
我们现在定义本章最重要的概念：图的拉普拉斯矩阵。实际上，正如我们将看到的，它有几种口味。

18.2. LAPLACIAN MATRICES OF GRAPHS  
18.2。图的拉普拉斯矩阵

## 18.2 Laplacian Matrices of Graphs 18.2图的拉普拉斯矩阵

Let us begin with directed graphs, although as we will see, graph Laplacians are fundamentally associated with undirected graph. The key proposition below shows how given an undirected graph G, for any orientation σ of G, Bσ(Bσ)> relates to the adjacency matrix A (where Bσ is the incidence matrix of the directed graph Gσ). We reproduce the proof in Gallier [72] (see also Godsil and Royle [77]).  
让我们从有向图开始，尽管正如我们将看到的，图拉普拉斯人基本上与无向图联系在一起。下面的关键命题展示了给定的无向图g，对于g的任何方向，bσ（bσ）>如何与邻接矩阵a（其中bσ是有向图gσ的关联矩阵）相关。我们用Gallier[72]复制了证据（另见Godsil和Royle[77]）。

Proposition 18.2. Given any undirected graph G, for any orientation σ of G, if Bσis the incidence matrix of the directed graph Gσ, A is the adjacency matrix of Gσ, and D is the degree matrix such that Dii = d(vi), then  
提案18.2.给定任意无向图g，对于g的任何方向σ，如果bσ是有向图gσ的关联矩阵，a是gσ的邻接矩阵，d是度矩阵，因此dii=d（vi），那么

Bσ(Bσ)> = D − A.  
bσ（bσ）>=d−a。

Consequently, L = Bσ(Bσ)> is independent of the orientation σ of G, and D−A is symmetric and positive semidefinite; that is, the eigenvalues of D − A are real and nonnegative. Proof. The entry Bσ(Bσ)>ij is the inner product of the ith row bσi , and the jth row bσj of Bσ.  
因此，l=bσ（bσ）>与g的方向σ无关，d−a是对称的正半定的，即d−a的特征值是实的和非负的。证据。条目bσ（bσ）>i j是第i行bσi的内积，bσj的第j行bσj。

|  |  |
| --- | --- |
| If i = j, then as 网络错误 |  |
|  网络错误  +1 网络错误   网络错误  bσik = −1 网络错误  0 网络错误 | if s(ek) = vi if t(ek) = vi otherwise 网络错误 |

we see that bσi · bσi = d(vi). If i =6 j, then bσi · bjσ = 06 iff there is some edge ek with s(ek) = vi and t(ek) = vj or vice-versa (which are mutually exclusive cases, since Gσ arises by orienting an undirected graph), in which case, bσi · bσj = −1. Therefore,  
我们看到bσi·bσi=d（vi）。如果i=6 J，那么bσi·b jσ=06如果有一些边Ek，s（Ek）=vi和t（Ek）=vj，反之亦然（这是互斥的情况，因为gσ是通过定向无向图产生的），在这种情况下，bσi·bσj=−1。因此，

Bσ(Bσ)> = D − A,  
bσ（bσ）>=d−a，

as claimed.  
如要求。

For every x ∈ Rm, we have  
对于每一个x∈rm，我们有

,  
，

since the Euclidean norm k k2 is positive (definite). Therefore, L = Bσ(Bσ)> is positive semidefinite. It is well-known that a real symmetric matrix is positive semidefinite iff its eigenvalues are nonnegative.   
因为欧几里得范数k k2是正的（确定的）。因此，l=bσ（bσ）>是半正定的。众所周知，实对称矩阵是半正定的，而其特征值是非负的。

Definition 18.15. The matrix L = Bσ(Bσ)> = D − A is called the (unnormalized) graph  
定义18.15.矩阵l=bσ（bσ）>=d−a称为（非正规化）图。

Laplacian of the graph Gσ. The (unnormalized) graph Laplacian of an undirected graph  
图Gσ的拉普拉斯函数。无向图的拉普拉斯图

G = (V,E) is defined by  
g=（v，e）的定义如下：

L = D − A.  
L=D−A。

For example, the graph Laplacian of graph G1 is  
例如，图g1的拉普拉斯图是

.  
.

Observe that each row of L sums to zero (because (Bσ)>1 = 0). Consequently, the vector 1 is in the nullspace of L.  
观察每行L的总和为零（因为（bσ）>1=0）。因此，向量1在l的空空间中。

Remarks:  
评论：

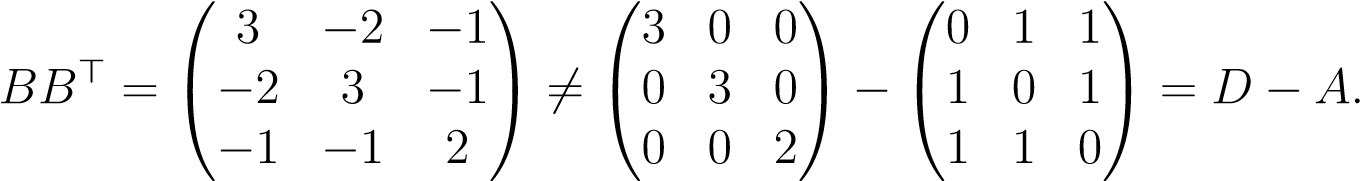
1. With the unoriented version of the incidence matrix (see Definition 18.8), it can be shown that  
   对于无定向版本的发病率矩阵（见定义18.8），可以证明：

BB> = D + A.  
bb>=d+a。

1. As pointed out by Evangelos Chatzipantazis, Proposition 18.2 in which the incidence matrix Bσ is replaced by the incidence matrix B of any arbitrary directed graph G does not hold. The problem is that such graphs may have both edges (vi,vj) and (vj,vi) between two distinct nodes vi and vj, and as a consequence, the inner product bi · bj = −2 instead of −1. A simple counterexample is given by the directed graph with three vertices and four edges whose incidence matrix is given by  
   正如Evangelos Chatzipantazis所指出的那样，18.2命题中，任何任意有向图G的关联矩阵b代替了关联矩阵b。问题是这样的图可能在两个不同的节点vi和vj之间都有边（vi，vj）和（vj，vi），因此内部积bi·bj=-2而不是-1。给出了三顶点四边有向图的一个简单反例，其关联矩阵由

.  
.

We have  
我们有



The natural generalization of the notion of graph Laplacian to weighted graphs is this:  
图拉普拉斯的概念对加权图的自然概括是：

Definition 18.16. Given any weighted graph G = (V,W) with V = {v1,...,vm}, the (unnormalized) graph Laplacian L(G) of G is defined by  
定义18.16.给定任意加权图g=（v，w），其中v=v1，…，vm，g的（非正规化）图拉普拉斯L（g）定义为

L(G) = D(G) − W,  
L（g）=D（g）−W，

where D(G) = diag(d1,...,dm) is the degree matrix of G (a diagonal matrix), with  
其中d（g）=diag（d1，…，dm）是g（对角矩阵）的度数矩阵，其中

.  
.

As usual, unless confusion arises, we write D instead of D(G) and L instead of L(G).  
像往常一样，除非出现混淆，我们写D而不是D（G），L而不是L（G）。

### 18.2. LAPLACIAN MATRICES OF GRAPHS 18.2。图的拉普拉斯矩阵

The graph Laplacian can be interpreted as a linear map from RV to itself. For all x ∈ RV , we have  
拉普拉斯图可以解释为从RV到自身的线性映射。对于所有的x∈rv，我们有

.  
.

It is clear from the equation L = D − W that each row of L sums to 0, so the vector 1 is the nullspace of L, but it is less obvious that L is positive semidefinite. One way to prove it is to generalize slightly the notion of incidence matrix.  
从方程l=d−w可以清楚地看出，每行l的总和为0，因此向量1是l的空空间，但不太明显l是半定的。证明这一点的一种方法是稍微概括一下关联矩阵的概念。

Definition 18.17. Given a weighted graph G = (V,W), with V = {v1,...,vm}, if {e1,..., en} are the edges of the underlying graph of G (recall that {vi,vj} is an edge of this graph iff wij > 0), for any oriented graph Gσ obtained by giving an orientation to the underlying graph of G, the incidence matrix Bσ of Gσ is the m × n matrix whose entries bij are given by  
定义18.17.给定一个加权图g=（v，w），其中v=v1，…，vm，如果e1，…，en是g的下一个图的边（回想一下v i，vj是这个图的边iff wij>0），对于通过给g的下一个图一个方向而得到的任何定向图gσ，gσi的关联矩阵bσs m×n矩阵，其条目bij由下式给出。

.  
.

For example, given the weight matrix  
例如，给定权重矩阵

,  
，

the incidence matrix B corresponding to the orientation of the underlying graph of W where an edge (i,j) is oriented positively iff i < j is  
当边缘（i，j）为正方向且i<j为正方向时，对应于基础图w方向的入射矩阵b

.  
.

The reader should verify that BB> = D − W. This is true in general, see Proposition 18.3.  
读者应确认bb>=d−w。这通常是正确的，见第18.3条建议。

It is easy to see that Proposition 18.1 applies to the underlying graph of G. For any oriented graph Gσ obtained from the underlying graph of G, the rank of the incidence matrix Bσ is equal to m−c, where c is the number of connected components of the underlying graph of G, and we have (Bσ)>1 = 0. We also have the following version of Proposition 18.2 whose proof is immediately adapted.  
很容易看出，命题18.1适用于g的下垫图。对于从g的下垫图获得的任何定向图gσ，关联矩阵bσ的秩等于m−c，其中c是g下垫图的连通分量的数目，我们得到（bσ）>1=0。我们也有以下版本的18.2号提案，其证据立即被采纳。

Proposition 18.3. Given any weighted graph G = (V,W) with V = {v1,...,vm}, if Bσ is the incidence matrix of any oriented graph Gσ obtained from the underlying graph of G and D is the degree matrix of G, then  
提案18.3.给定任意加权图g=（v，w），其中v=v1，…，vm，如果bσ是从g的底层图中得到的任意定向图g的关联矩阵，d是g的度矩阵，那么

Bσ(Bσ)> = D − W = L.  
bσ（bσ）>=d−w=l。

Consequently, Bσ(Bσ)> is independent of the orientation of the underlying graph of G and L = D − W is symmetric and positive semidefinite; that is, the eigenvalues of L = D − W are real and nonnegative.  
因此，bσ（bσ）>与g的下垫图的方向无关，l=d−w是对称的正半定的，也就是说，l=d−w的特征值是实的和非负的。

Another way to prove that L is positive semidefinite is to evaluate the quadratic form x>Lx.  
证明l是半正定的另一种方法是求二次型x>lx。

Proposition 18.4. For any m × m symmetric matrix W = (wij), if we let L = D − W where D is the degree matrix associated with W (that is, ), then we have  
提案18.4.对于任何m×m对称矩阵w=（wij），如果我们让l=d−w，其中d是与w相关联的度数矩阵（即，），那么我们有

for all x ∈ Rm.  
对于所有x∈rm。

Consequently, x>Lx does not depend on the diagonal entries in W, and if wij ≥ 0 for all i,j ∈ {1,...,m}, then L is positive semidefinite.  
因此，x>lx不依赖于w中的对角线项，如果所有i，j∈1，…，m的wij≥0，则l为正半定。

Proof. We have  
证据。我们有

!  
！

Obviously, the quantity on the right-hand side does not depend on the diagonal entries in W, and if wij ≥ 0 for all i,j, then this quantity is nonnegative.   
显然，右边的数量不依赖于w中的对角线项，如果w i j对于所有i，j都大于等于0，那么这个数量是非负的。

Proposition 18.4 immediately implies the following facts: For any weighted graph G = (V,W),  
命题18.4立即暗示了以下事实：对于任何加权图g=（v，w），

1. The eigenvalues 0 = λ1 ≤ λ2 ≤ ... ≤ λm of L are real and nonnegative, and there is an orthonormal basis of eigenvectors of L.  
   特征值0=λ1≤λ2≤…L的λm是实的和非负的，并且L的特征向量有一个正交基。
2. The smallest eigenvalue λ1 of L is equal to 0, and 1 is a corresponding eigenvector.  
   L的最小特征值λ1等于0，1是对应的特征向量。

It turns out that the dimension of the nullspace of L (the eigenspace of 0) is equal to the number of connected components of the underlying graph of G.  
结果表明，L的零空间维数（特征空间为0）等于G的底层图的连通分量的个数。

### 18.3. NORMALIZED LAPLACIAN MATRICES OF GRAPHS 18.3。图的正规拉普拉斯矩阵

Proposition 18.5. Let G = (V,W) be a weighted graph. The number c of connected components K1,...,Kc of the underlying graph of G is equal to the dimension of the nullspace of L, which is equal to the multiplicity of the eigenvalue 0. Furthermore, the nullspace of L has a basis consisting of indicator vectors of the connected components of G, that is, vectors (f1,...,fm) such that fj = 1 iff vj ∈ Ki and fj = 0 otherwise.  
提案18.5。设g=（v，w）为加权图。G下垫图的连通分量k1，…，kc的个数c等于L的零空间的维数，等于特征值0的重数。此外，L的零空间有一个基，由G的连通分量的指示向量构成，即向量（f1，…，fm），使fj=1 iff vj∈ki，否则fj=0。

Proof. Since L = BB> for the incidence matrix B associated with any oriented graph obtained from G, and since L and B> have the same nullspace, by Proposition 18.1, the dimension of the nullspace of L is equal to the number c of connected components of G and the indicator vectors of the connected components of G form a basis of Ker(L).   
证据。由于l=b b>对于与从g得到的任何定向图相关联的关联矩阵b，并且l和b>具有相同的零空间，根据命题18.1，l的零空间的维数等于g的连接分量的C数和连接分量的指示向量。t的g构成了ker（l）的基础。

Proposition 18.5 implies that if the underlying graph of G is connected, then the second eigenvalue λ2 of L is strictly positive.  
命题18.5意味着，如果G的下垫图是连通的，那么L的第二特征值λ2是严格正的。

Remarkably, the eigenvalue λ2 contains a lot of information about the graph G (assuming that G = (V,E) is an undirected graph). This was first discovered by Fiedler in 1973, and for this reason, λ2 is often referred to as the Fiedler number. For more on the properties of the Fiedler number, see Godsil and Royle [77] (Chapter 13) and Chung [39]. More generally, the spectrum (0,λ2,...,λm) of L contains a lot of information about the combinatorial structure of the graph G. Leverage of this information is the object of spectral graph theory.  
值得注意的是，特征值λ2包含了大量关于图G的信息（假设g=（v，e）是无向图）。这是费德勒在1973年首次发现的，因此，λ2通常被称为费德勒数。关于费德勒数属性的更多信息，见Godsil和Royle[77]（第13章）和Chung[39]。一般来说，L的谱（0，λ2，…，λm）包含了大量关于图G组合结构的信息，利用这些信息是谱图理论的研究对象。

## 18.3 Normalized Laplacian Matrices of Graphs 18.3图的正规拉普拉斯矩阵

It turns out that normalized variants of the graph Laplacian are needed, especially in applications to graph clustering. These variants make sense only if G has no isolated vertices.  
结果表明，需要图形拉普拉斯的规范化变量，特别是在图形聚类的应用中。只有当g没有孤立的顶点时，这些变量才有意义。

Definition 18.18. Given a weighted graph G = (V,W), a vertex u ∈ V is isolated if it is not incident to any other vertex. This means that every row of W contains some strictly positive entry.  
定义18.18.给定一个加权图g=（v，w），如果一个顶点u∈v不与任何其他顶点相关联，则它是孤立的。这意味着W的每一行都包含一些严格的正数。

If G has no isolated vertices, then the degree matrix D contains positive entries, so it is invertible and D−1/2 makes sense; namely  
如果g没有孤立的顶点，那么度数矩阵d包含正项，因此它是可逆的，d−1/2是有意义的；即

D−1/2 = diag(,  
D−1/2=诊断（，

and similarly for any real exponent α.  
与任何实数指数α类似。

Definition 18.19. Given any weighted directed graph G = (V,W) with no isolated vertex and with V = {v1,...,vm}, the (normalized) graph Laplacians Lsym and Lrw of G are defined by  
定义18.19.给定任何无孤立顶点的加权有向图g=（v，w），以及v=v1，…，vm，（归一化）图g的拉普拉斯LSYM和lrw定义为

Lsym = D−1/2LD−1/2 = I − D−1/2WD−1/2 Lrw = D−1L = I − D−1W.  
LSYM=D−1/2LD−1/2=I−D−1/2WD−1/2 LRW=D−1L=I−D−1W。

Observe that the Laplacian Lsym = D−1/2LD−1/2 is a symmetric matrix (because L and  
观察拉普拉斯LSYM=d−1/2ld−1/2是一个对称矩阵（因为l和

D−1/2 are symmetric) and that  
d−1/2是对称的）和

Lrw = D−1/2LsymD1/2.  
lrw=d−1/2lsymd1/2。

The reason for the notation Lrw is that this matrix is closely related to a random walk on the graph G.  
符号lrw的原因是这个矩阵与图g上的随机游动密切相关。

Example 18.1. As an example, the matrices Lsym and Lrw associated with the graph G1  
例18.1。例如，与图g1相关的矩阵lsym和lrw

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| are 网络错误 |  |  |  |  |  |
| and 网络错误 |  | −10.0000.3536 网络错误  −00..28872887 −0.3536 网络错误  − 网络错误 | −00..40822887 网络错误  −1.0000 网络错误  −0.03333 网络错误 | 0 网络错误  −00..28873333 网络错误  −1.0000 网络错误  −0.4082 网络错误 | 0  网络错误  −0.3536 网络错误  0  网络错误  −0.4082 网络错误  1.0000 网络错误 |
|  |  | −10.0000.5000 网络错误  −00..33333333 −0.5000 网络错误  − 网络错误 | −00..50002500 网络错误  −1.0000 网络错误  −0.03333 网络错误 | 0 网络错误  −00..25003333 网络错误  −1.0000 网络错误  −0.5000 网络错误 | 0  网络错误  −0.2500 网络错误  0 . 网络错误  −0.3333 网络错误  1.0000 网络错误 |

Since the unnormalized Laplacian L can be written as L = BB>, where B is the incidence matrix of any oriented graph obtained from the underlying graph of G = (V,W), if we let  
由于非正规拉普拉斯L可以写成l=b b>，其中b是从基础图g=（v，w）得到的任何定向图的关联矩阵，如果我们

Bsym = D−1/2B,  
Bsym=d−1/2b，

we get  
我们得到

.  
.

In particular, for any singular decomposition Bsym = UΣV > of Bsym (with U an m × m orthogonal matrix, Σ a “diagonal” m×n matrix of singular values, and V an n×n orthogonal matrix), the eigenvalues of Lsym are the squares of the top m singular values of Bsym, and the vectors in U are orthonormal eigenvectors of Lsym with respect to these eigenvalues (the squares of the top m diagonal entries of Σ). Computing the SVD of Bsym generally yields more accurate results than diagonalizing Lsym, especially when Lsym has eigenvalues with high multiplicity.  
特别是对于任意一个bsym的奇异分解bsym=u∑v>（有u个m×m正交矩阵，∑一个“对角”m×n奇异值矩阵，v个n×n正交矩阵），lsym的特征值是bsym上m奇异值的平方，u中的向量是正交的。关于这些特征值的lsym的法向特征向量（∑的上m个对角线项的平方）。计算bsym的SVD通常比对角化lsym得到更精确的结果，特别是当lsym具有高重特征值时。

There are simple relationships between the eigenvalues and the eigenvectors of Lsym, and Lrw. There is also a simple relationship with the generalized eigenvalue problem Lx = λDx.  
LSYM和LRW的特征值与特征向量之间存在简单的关系。与广义特征值问题lx=λdx也有一个简单的关系。

Proposition 18.6. Let G = (V,W) be a weighted graph without isolated vertices. The graph Laplacians, L,Lsym, and Lrw satisfy the following properties:  
提案18.6.设g=（v，w）为无孤立顶点的加权图。拉普拉斯图、L、Lsym和Lrw满足以下特性：

### 18.3. NORMALIZED LAPLACIAN MATRICES OF GRAPHS 18.3。图的正规拉普拉斯矩阵

1. The matrix Lsym is symmetric and positive semidefinite. In fact,  
   矩阵LSYM是对称的半正定矩阵。事实上，

for all x ∈ Rm.  
对于所有x∈rm。

1. The normalized graph Laplacians Lsym and Lrw have the same spectrum  
   归一化图拉普拉斯函数lsym和lrw具有相同的谱。

(0 = ν1 ≤ ν2 ≤ ... ≤ νm), and a vector u 6 = 0is an eigenvector of Lrw for λ iff D/2u is an eigenvector of Lsym for λ.  
（0=霏1≤霏2≤…且向量u 6=0是λiff d/2u的lrw的特征向量是λ的lsym的特征向量。

1. The graph Laplacians L and Lsym are symmetric and positive semidefinite.  
   图拉普拉斯l和lsym是对称的、半正定的。
2. A vector u = 06 is a solution of the generalized eigenvalue problem Lu = λDu iff D1/2u is an eigenvector of Lsym for the eigenvalue λ iff u is an eigenvector of Lrw for the eigenvalue λ.  
   向量u=06是广义特征值问题lu=λdu iff d1/2u是特征值λ的LSYM特征向量，iff u是特征值λ的LRW特征向量。
3. The graph Laplacians, L and Lrw have the same nullspace. For any vector u, we have u ∈ Ker(L) iff D1/2u ∈ Ker(Lsym).  
   图拉普拉斯，l和lrw有相同的空空间。对于任何向量u，我们都有u∈ker（l）iff d1/2u∈ker（lsym）。
4. The vector 1 is in the nullspace of Lrw, and D1/21 is in the nullspace of Lsym.  
   矢量1在lrw的空空间中，d1/21在lsym的空空间中。
5. For every eigenvalue νi of the normalized graph Laplacian Lsym, we have 0 ≤ νi ≤ 2. Furthermore, νm = 2 iff the underlying graph of G contains a nontrivial connected bipartite component.  
   对于归一化图拉普拉斯LSYM的每一个特征值，我们有0≤νi≤2。此外，如果g的下垫图包含一个非平凡的连接二部分量，则v m=2。
6. If m ≥ 2 and if the underlying graph of G is not a complete graph,1 then ν2 ≤ 1.  
   如果m≥2，如果g的下垫图不是一个完整的图，1，则v 2≤1。

Furthermore the underlying graph of G is a complete graph iff.  
此外，G的底层图是一个完整的IFF图。

1. If m ≥ 2 and if the underlying graph of G is connected, then ν2 > 0.  
   如果m≥2且G的下垫图连通，则v 2>0。
2. If m ≥ 2 and if the underlying graph of G has no isolated vertices, then .  
   如果m≥2，并且g的基础图没有孤立的顶点，那么。

Proof. (1) We have Lsym = D−1/2LD−1/2, and D−1/2 is a symmetric invertible matrix (since it is an invertible diagonal matrix). It is a well-known fact of linear algebra that if B is an invertible matrix, then a matrix S is symmetric, positive semidefinite iff BSB> is symmetric, positive semidefinite. Since L is symmetric, positive semidefinite, so is Lsym = D−1/2LD−1/2. The formula  
证据。（1）我们有lsym=d−1/2ld−1/2，d−1/2是对称可逆矩阵（因为它是可逆对角矩阵）。线性代数的一个众所周知的事实是，如果b是可逆矩阵，那么矩阵s是对称的，半正定的iff bsb>是对称的，半正定的。因为l是对称的，半正定的，所以lsym=d−1/2ld−1/2也是。公式

for all x ∈ Rm  
对于所有x∈rm

follows immediately from Proposition 18.4 by replacing x by D−1/2x, and also shows that Lsym is positive semidefinite.  
紧随命题18.4，用d−1/2x替换x，还表明lsym是正半定的。

* 1. Since  
     自从

Lrw = D−1/2LsymD1/2,  
lrw=d−1/2lsymd1/2，

the matrices Lsym and Lrw are similar, which implies that they have the same spectrum. In fact, since D1/2 is invertible,  
矩阵lsym和lrw相似，这意味着它们具有相同的光谱。实际上，由于d1/2是可逆的，

Lrwu = D−1Lu = λu  
LrWu=d−1lu=λu

iff  
敌我识别

D−1/2Lu = λD1/2u  
D−1/2lu=λd 1/2u

iff  
敌我识别

D−1/2LD−1/2D1/2u = LsymD1/2u = λD1/2u,  
d−1/2ld−1/2d 1/2u=lsymd1/2u=λd1/2u，

which shows that a vector u = 06 is an eigenvector of Lrw for λ iff D1/2u is an eigenvector of Lsym for λ.  
结果表明，对于λiff d1/2u，向量u=06是lrw的特征向量，对于λ，向量l=06是lsym的特征向量。

* 1. We already know that L and Lsym are positive semidefinite. (4) Since D−1/2 is invertible, we have  
     我们已经知道l和lsym是正半定的。（4）由于d−1/2是可逆的，我们有

Lu = λDu  
Lu=λdu

iff  
敌我识别

D−1/2Lu = λD1/2u  
D−1/2lu=λd 1/2u

iff  
敌我识别

D−1/2LD−1/2D1/2u = LsymD1/2u = λD1/2u,  
d−1/2ld−1/2d 1/2u=lsymd1/2u=λd1/2u，

which shows that a vector u = 06 is a solution of the generalized eigenvalue problem Lu = λDu iff D1/2u is an eigenvector of Lsym for the eigenvalue λ. The second part of the statement follows from (2).  
结果表明，向量u=06是广义特征值问题lu=λdu iff d1/2u的解，是特征值λ的Lsym的特征向量。陈述的第二部分来自（2）。

* 1. Since D−1 is invertible, we have Lu = 0 iff D−1Lu = Lrwu = 0. Similarly, since D−1/2 is invertible, we have Lu = 0 iff D−1/2LD−1/2D1/2u = 0 iff D1/2u ∈ Ker(Lsym).  
     因为d−1是可逆的，所以当d−1lu=lrwu=0时，我们得到lu=0。同样，由于d−1/2是可逆的，我们得到lu=0 iff d−1/2ld−1/2d1/2u=0 iff d1/2u∈ker（lsym）。
  2. Since L1 = 0, we get Lrw1 = D−1L1 = 0. That D1/21 is in the nullspace of Lsym follows from (2). Properties (7)–(10) are proven in Chung [39] (Chapter 1). The eigenvalues the matrices Lsym and Lrw from Example 18.1 are  
     因为l1=0，我们得到lrw1=d−1l1=0。d1/21在lsym的空白处，从（2）开始。性能（7）–（10）在Chung[39]中得到证明（第1章）。例18.1中矩阵lsym和lrw的特征值为

0, 7257, 1.1667, 1.5, 1.6076.  
0，7257，1.1667，1.5，1.6076。

On the other hand, the eigenvalues of the unormalized Laplacian for G1 are  
另一方面，非正规拉普拉斯的特征值为

0, 1.5858, 3, 4.4142, 5.  
0，1.5858，3，4.4142，5.

Remark: Observe that although the matrices Lsym and Lrw have the same spectrum, the matrix Lrw is generally not symmetric, whereas Lsym is symmetric.  
注：观察到尽管矩阵lsym和lrw具有相同的频谱，但矩阵lrw一般不对称，而lsym是对称的。

A version of Proposition 18.5 also holds for the graph Laplacians Lsym and Lrw. This follows easily from the fact that Proposition 18.1 applies to the underlying graph of a weighted graph. The proof is left as an exercise.  
命题18.5的一个版本也适用于拉普拉斯图lsym和lrw。这很容易从18.1命题适用于加权图的基础图这一事实得出。证据留作练习。

### 18.4. GRAPH CLUSTERING USING NORMALIZED CUTS 18.4。使用标准化切割的图形聚类

Proposition 18.7. Let G = (V,W) be a weighted graph. The number c of connected components K1,...,Kc of the underlying graph of G is equal to the dimension of the nullspace of both Lsym and Lrw, which is equal to the multiplicity of the eigenvalue 0. Furthermore, the nullspace of Lrw has a basis consisting of indicator vectors of the connected components of G, that is, vectors (f1,...,fm) such that fj = 1 iff vj ∈ Ki and fj = 0 otherwise. For Lsym, a basis of the nullpace is obtained by multiplying the above basis of the nullspace of Lrw by D1/2.  
提案18.7。设g=（v，w）为加权图。G下垫图的连通分量k1，…，kc的个数c等于Lsym和Lrw的空空间的维数，等于特征值0的重数。此外，LRW的零空间有一个由G的连通分量的指示向量构成的基，即向量（f1，…，fm），否则fj=1 iff vj∈ki，fj=0。对于lsym，空空间的基础是通过将lrw的空空间的上述基础乘以d1/2得到的。

A particularly interesting application of graph Laplacians is graph clustering.  
图拉普拉斯的一个特别有趣的应用是图聚类。

## 18.4 Graph Clustering Using Normalized Cuts 18.4使用标准化切割的图形聚类

In order to explain this problem we need some definitions.  
为了解释这个问题，我们需要一些定义。

Definition 18.20. Given any subset of nodes A ⊆ V , we define the volume vol(A) of A as the sum of the weights of all edges adjacent to nodes in A:  
定义18.20。给定节点a v的任何子集，我们将a的体积vol（a）定义为a中与节点相邻的所有边的权重之和：

m  
米

vol(A) = XXwij.  
体积（A）=xxwij。

vi∈A j=1  
vi∈a j=1

Given any two subsets A,B ⊆ V (not necessarily distinct), we define links(A,B) by  
对于任意两个子集a，b\_v（不一定是不同的），我们定义链接（a，b）的方式是

links(A,B) = X wij.  
链接（a，b）=x wij。

vi∈A,vj∈B  
vi∈a，vj∈b

The quantity links(A,A) = links(A,A) (where A = V − A denotes the complement of A in V ) measures how many links escape from A (and A). We define the cut of A as  
数量链接（A，A）=Links（A，A）（其中A=V−A表示V中A的补码）测量从A（和A）中漏出的链接数。我们将a的切割定义为

cut(A) = links(A,A).  
剪切（A）=Links（A，A）。

The notion of volume is illustrated in Figure 18.5 and the notions of cut is illustrated in Figure 18.6.  
体积的概念如图18.5所示，切割的概念如图18.6所示。

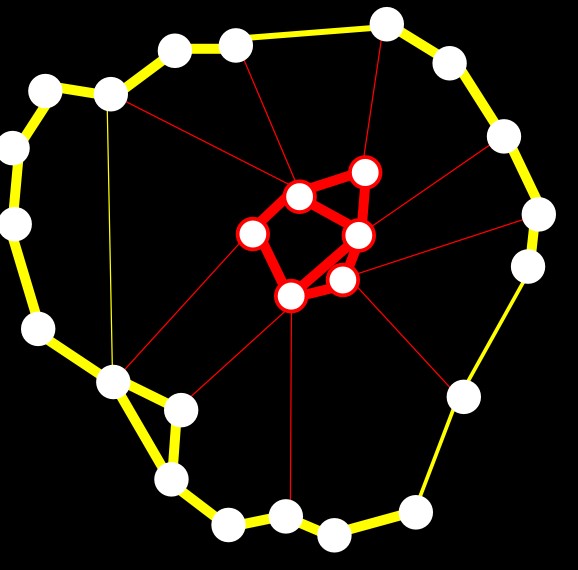


Figure 18.5: Volume of a set of nodes.  
图18.5：一组节点的体积。

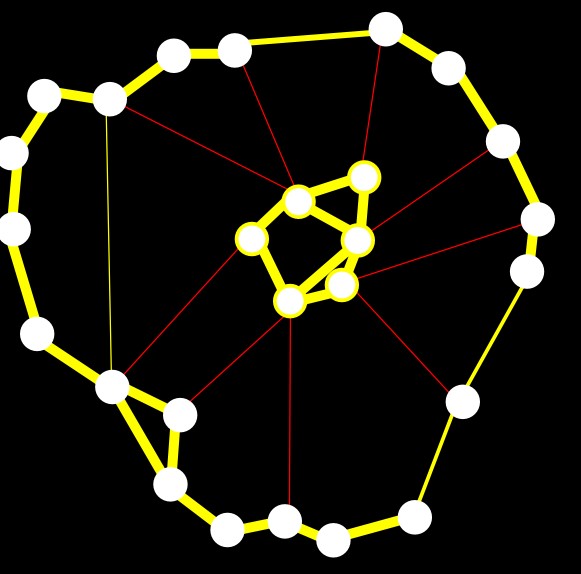


Figure 18.6: A cut involving the set of nodes in the center and the nodes on the perimeter.  
图18.6：包括中心节点集和周界节点的切口。

The above concepts play a crucial role in the theory of normalized cuts. This beautiful and deeply original method first published in Shi and Malik [155], has now come to be a “textbook chapter” of computer vision and machine learning. It was invented by Jianbo Shi and Jitendra Malik and was the main topic of Shi’s dissertation. This method was extended to K ≥ 3 clusters by Stella Yu in her dissertation [185] and is also the subject of Yu and Shi  
上述概念在标准化切割理论中起着至关重要的作用。这一优美而深刻的原创方法首次发表在Shi和Malik[155]上，现已成为计算机视觉和机器学习的“教科书章节”。它是石建波和马立克共同发明的，是石建波博士论文的主题。该方法在论文[185]中被Stella Yu推广到k≥3簇，也是Yu和Shi的研究课题。

[187].  
〔187〕。

Given a set of data, the goal of clustering is to partition the data into different groups according to their similarities. When the data is given in terms of a similarity graph G, where the weight wij between two nodes vi and vj is a measure of similarity of vi and vj, the problem can be stated as follows: Find a partition (A1,...,AK) of the set of nodes V into different groups such that the edges between different groups have very low weight (which indicates that the points in different clusters are dissimilar), and the edges within a group have high weight (which indicates that points within the same cluster are similar).  
给定一组数据，聚类的目的是根据数据的相似性将其划分为不同的组。当用相似度图G给出数据时，其中两个节点vi和vj之间的权重wij是vi和vj相似度的度量，问题可以表述为：将一组节点v的分区（a1，…，ak）分成不同的组，这样不同的组之间的边UPS的权重非常低（表示不同集群中的点不同），而一组中的边缘权重较高（表示同一集群中的点相似）。

The above graph clustering problem can be formalized as an optimization problem, using the notion of cut mentioned earlier. If we want to partition V into K clusters, we can do so by finding a partition (A1,...,AK) that minimizes the quantity  
利用前面提到的割的概念，上述图聚类问题可以形式化为一个优化问题。如果我们想把v划分成k簇，我们可以通过找到一个最小化数量的分区（a1，…，ak）来实现。

K K  
K-K

cut(links(Ai,Ai).  
剪切（链接（ai，ai）。

=1 =1  
＝1＝1

For K = 2, the mincut problem is a classical problem that can be solved efficiently, but in practice, it does not yield satisfactory partitions. Indeed, in many cases, the mincut solution separates one vertex from the rest of the graph. What we need is to design our cost function in such a way that it keeps the subsets Ai “reasonably large” (reasonably balanced).  
对于k=2，mincut问题是一个可以有效解决的经典问题，但在实际应用中，它不能产生令人满意的分区。实际上，在许多情况下，mincut解将一个顶点与图中的其他顶点分开。我们需要的是设计我们的成本函数，使子集合ai“相当大”（相当平衡）。

An example of a weighted graph and a partition of its nodes into two clusters is shown in Figure 18.7.  
图18.7显示了一个加权图及其节点划分为两个集群的示例。

### 18.5. SUMMARY 18.5。总结

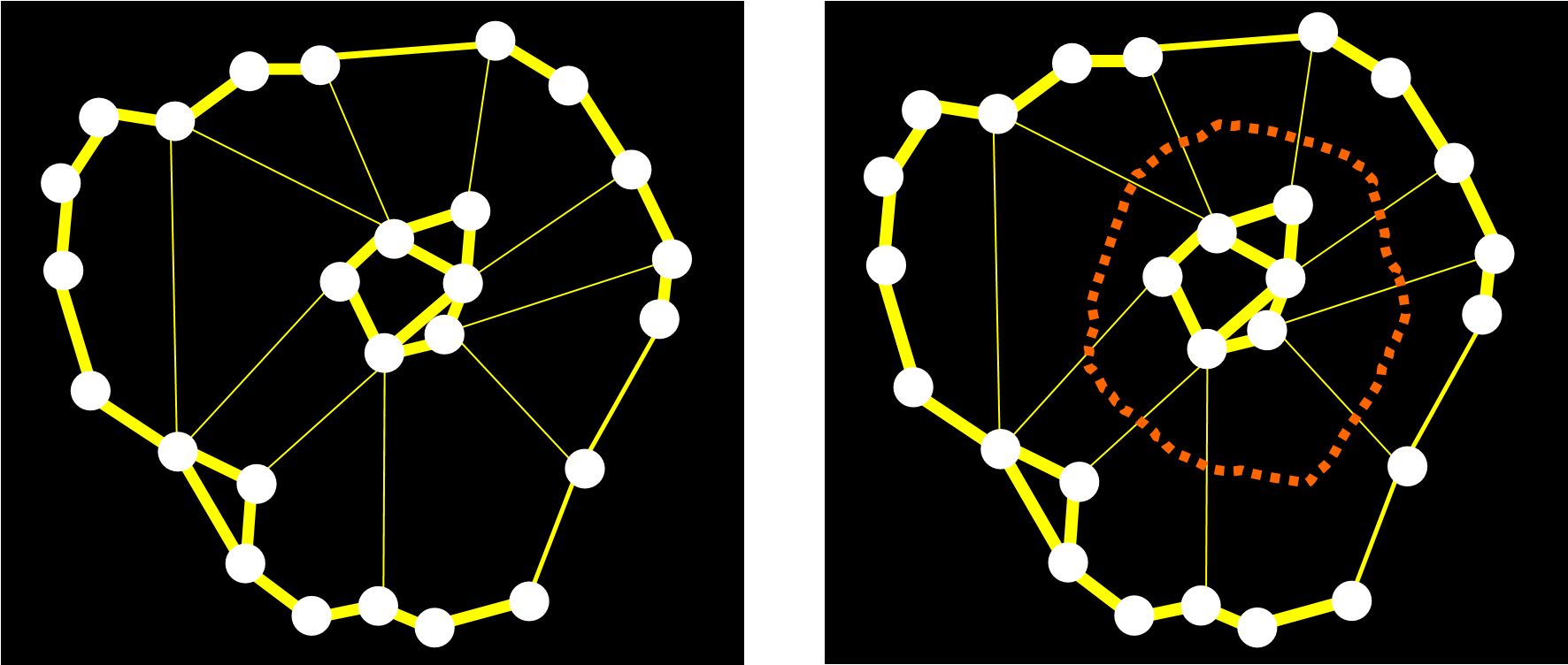


Figure 18.7: A weighted graph and its partition into two clusters.  
图18.7：加权图及其划分成两个簇。

A way to get around this problem is to normalize the cuts by dividing by some measure of each subset Ai. A solution using the volume vol(Ai) of Ai (for K = 2) was proposed and investigated in a seminal paper of Shi and Malik [155]. Subsequently, Yu (in her dissertation [185]) and Yu and Shi [187] extended the method to K > 2 clusters. The idea is to minimize the cost function  
解决这个问题的一种方法是通过除以每个子集ai的某个度量来规范化切割。在Shi和Malik的一篇开创性论文[155]中，提出并研究了一种利用ai体积（k=2）的解决方案。随后，Yu（在她的论文[185]中）和Yu和Shi[187]将该方法扩展到k>2个簇。其思想是最小化成本函数

KK  
KK

Ncut(A1,...,AK) = X links(Ai,Ai) = X cut(Ai,Ai). i=1 vol(Ai) i=1 vol(Ai)  
ncut（a1，…，ak）=x链接（ai，ai）=x剪切（ai，ai）。i=1卷（ai）i=1卷（ai）

The next step is to express our optimization problem in matrix form, and this can be done in terms of Rayleigh ratios involving the graph Laplacian in the numerators. This theory is very beautiful, but we do not have the space to present it here. The interested reader is referred to Gallier [70].  
下一步是用矩阵形式表示我们的优化问题，这可以用分子中含有拉普拉斯图的瑞利比来表示。这个理论很美，但是我们没有空间来展示它。感兴趣的读者可参考Gallier[70]。

## 18.5 Summary 18.5总结

The main concepts and results of this chapter are listed below:  
本章的主要概念和结果如下：

* Directed graphs, undirected graphs.  
  有向图，无向图。
* Incidence matrices, adjacency matrices.  
  关联矩阵，邻接矩阵。
* Weighted graphs.  
  加权图。
* Degree matrix.  
  度矩阵。
* Graph Laplacian (unnormalized).  
  拉普拉斯图（非标准化）。
* Normalized graph Laplacian.  
  规范化拉普拉斯图。
* Spectral graph theory.  
  光谱图理论。
* Graph clustering using normalized cuts.  
  使用标准化切割进行图形聚类。

## 18.6 Problems 18.6问题

Problem 18.1. Find the unnormalized Laplacian of the graph representing a triangle and of the graph representing a square.  
问题18.1。求表示三角形的图和表示正方形的图的非正规拉普拉斯。

Problem 18.2. Consider the complete graph Km on m ≥ 2 nodes.  
问题18.2。考虑m≥2个节点上的完整图km。

1. Prove that the normalized Laplacian Lsym of K is  
   证明K的正规拉普拉斯LSYM是

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  网络错误  1 网络错误  −1/(m − 1) Lsym =  ... 网络错误   网络错误   网络错误  −1/(m − 1) 网络错误   网络错误  −1/(m − 1) 网络错误 | −1/(m1 − 1) 网络错误  ... 网络错误  −11//((mm −− 1)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 | −11//((.mm..−− 1)1) 网络错误  − 网络错误  1 网络错误  −1/(m − 1) 网络错误 | −1/(m − 1) 网络错误  −1/(m − 1) 网络错误  ... . 网络错误  −1/(m − 1) 1 网络错误 |

1. Prove that the characteristic polynomial of Lsym is  
   证明了LSYM的特征多项式是

.  
.

Hint. First subtract the second column from the first, factor λ − m/(m − 1), and then add the first row to the second. Repeat this process. You will end up with the determinant  
暗示。首先从第一列中减去第二列，系数λ−m/（m−1），然后将第一行添加到第二行。重复此过程。你最终会得到行列式

.  
.

Problem 18.3. Consider the complete bipartite graph Km,n on m + n ≥ 3 nodes, with edges between each of the first m ≥ 1 nodes to each of the last n ≥ 1 nodes. Prove that the eigenvalues of the normalized Laplacian Lsym of Km,n are 0 with multiplicity m + n − 2 and 1 with multiplicity 2.  
问题18.3。考虑完整的二部图km，n在m+n≥3个节点上，第一个m≥1个节点与最后n≥1个节点之间的边。证明了m+n-2的归一化拉普拉斯LSYM的特征值为0，2的归一化拉普拉斯LSYM的特征值为1。

Problem 18.4. Let G be a graph with a set of nodes V with m ≥ 2 elements, without isolated nodes, and let Lsym = D−1/2LD−1/2 be its normalized Laplacian (with L its unnormalized Laplacian).  
问题18.4.设g为一组节点v的图，其中m≥2个元素，无孤立节点；设lsym=d−1/2ld−1/2为其归一化拉普拉斯函数（L为其非归一化拉普拉斯函数）。

### 18.6. PROBLEMS 18.6。问题

1. For any y ∈ RV , consider the Rayleigh ratio  
   对于任意y∈rv，考虑瑞利比

.  
.

Prove that if x = D−1/2y, then  
证明如果x=d−1/2y，那么

.  
.

1. Prove that the second eigenvalue ν2 of Lsym is given by  
   证明lsym的第二特征值v 2由下式给出

.  
.

1. Prove that the largest eigenvalue νm of Lsym is given by  
   证明了lsym的最大特征值νm由下式给出

.  
.

Problem 18.5. Let G be a graph with a set of nodes V with m ≥ 2 elements, without isolated nodes. If 0 = ν1 ≤ ν1 ≤ ... ≤ νm are the eigenvalues of Lsym, prove the following properties:  
问题18.5。设g为一组节点v的图，其中m≥2个元素，没有孤立的节点。如果0=ν1≤ν1≤…≤νm是lsym的特征值，证明了以下性质：

1. We have ν1 + ν2 + ··· + νm = m.  
   我们得到了，θ1+θ2+·········，θm=m。
2. We have ν2 ≤ m/(m−1), with equality holding iff G = Km, the complete graph on m nodes.  
   我们得到了θ2≤m/（m-1），等式为iff g=km，m节点上的完整图。
3. We have νm ≥ m/(m − 1).  
   我们有v m≥m/（m-1）。
4. If G is not a complete graph, then ν2 ≤ 1  
   如果g不是一个完整的图，那么v 2≤1

Hint. If a and b are nonadjacent nodes, consider the function x given by  
暗示。如果a和b是非相邻节点，则考虑由

|  |  |
| --- | --- |
|  | if v = a if v = b 网络错误  if v =6 a,b, 网络错误 |

and use Problem 18.4(2).  
使用问题18.4（2）。

1. Prove that νm ≤ 2. Prove that νm = 2 iff the underlying graph of G contains a nontrivial connected bipartite component.  
   证明νm≤2。证明了当g的底层图包含一个非平凡的连通二部分量时，νm=2。

Hint. Use Problem 18.4(3).  
暗示。使用问题18.4（3）。

1. Prove that if G is connected, then ν2 > 0.  
   证明如果G连通，那么v 2>0。

Problem 18.6. Let G be a graph with a set of nodes V with m ≥ 2 elements, without isolated nodes. Let vol(G) = Pv∈V dv and let  
问题18.6.设g为一组节点v的图，其中m≥2个元素，没有孤立的节点。设vol（g）=pv∈v dv，设

.  
.

Prove that  
证明这一点

.  
.

Problem 18.7. Let G be a connected bipartite graph. Prove that if ν is an eigenvalue of Lsym, then 2 − ν is also an eigenvalue of Lsym.  
问题18.7。设G为连通二部图。证明如果nv是lsym的特征值，那么2−nv也是lsym的特征值。

Problem 18.8. Prove Proposition 18.7.  
问题18.8。证明18.7号提案。

Chapter 19  
第十九章

# Spectral Graph Drawing 光谱图绘制

## 19.1 Graph Drawing and Energy Minimization 19.1图形绘制和能量最小化

Let G = (V,E) be some undirected graph. It is often desirable to draw a graph, usually in the plane but possibly in 3D, and it turns out that the graph Laplacian can be used to design surprisingly good methods. Say |V | = m. The idea is to assign a point ρ(vi) in Rn to the vertex vi ∈ V , for every vi ∈ V , and to draw a line segment between the points ρ(vi) and ρ(vj) iff there is an edge {vi,vj}.  
设g=（v，e）为无向图。通常需要绘制一个图形，通常是在平面上，但可能是在三维中，结果证明，拉普拉斯图可以用来设计出奇的好方法。假设v=m，其思想是将RN中的点ρ（vi）赋给顶点vi∈v，对于每个vi∈v，并在点ρ（vi）和ρ（vj）之间画一条直线段，如果有边vi，vj。

Definition 19.1. Let G = (V,E) be some undirected graph with m vertices. A graph drawing is a function ρ: V → Rn, for some n ≥ 1. The matrix of a graph drawing ρ (in Rn) is a m × n matrix R whose ith row consists of the row vector ρ(vi) corresponding to the point representing vi in Rn.  
定义19.1.设g=（v，e）为具有m个顶点的无向图。图形绘制是函数ρ：v→rn，对于某些n≥1。图的矩阵ρ（在r n中）是m×n矩阵r，其第i行由行向量ρ（vi）组成，与在rn中表示vi的点相对应。

For a graph drawing to be useful we want n ≤ m; in fact n should be much smaller than m, typically n = 2 or n = 3.  
为了使图形绘图有用，我们希望n≤m；实际上n应该比m小得多，通常n=2或n=3。

Definition 19.2. A graph drawing is balanced iff the sum of the entries of every column of the matrix of the graph drawing is zero, that is,  
定义19.2.图的绘制是平衡的，如果图的矩阵的每一列的条目之和为零，也就是说，

1>R = 0.  
1>r=0。

If a graph drawing is not balanced, it can be made balanced by a suitable translation. We may also assume that the columns of R are linearly independent, since any basis of the column space also determines the drawing. Thus, from now on, we may assume that n ≤ m.  
如果图的绘制不平衡，可以通过适当的翻译使其平衡。我们也可以假设r的列是线性独立的，因为列空间的任何基础也决定了绘图。因此，从现在开始，我们可以假定n≤m。

Remark: A graph drawing ρ: V → Rn is not required to be injective, which may result in degenerate drawings where distinct vertices are drawn as the same point. For this reason, we prefer not to use the terminology graph embedding, which is often used in the literature. This is because in differential geometry, an embedding always refers to an injective map. The term graph immersion would be more appropriate.  
注：图的ρ：v→rn不需要求内射，这可能导致退化图，其中不同的顶点绘制为同一点。因此，我们不喜欢使用文献中经常使用的术语图嵌入。这是因为在微分几何中，嵌入总是指一个内射映射。术语图浸入更合适。

603  
六百零三

As explained in Godsil and Royle [77], we can imagine building a physical model of G by connecting adjacent vertices (in Rn) by identical springs. Then it is natural to consider a representation to be better if it requires the springs to be less extended. We can formalize this by defining the energy of a drawing R by  
如Godsil和Royle[77]所述，我们可以想象通过相同的弹簧连接相邻顶点（RN）来建立G的物理模型。那么，如果要求弹簧的延伸度较小，那么考虑更好的表示是很自然的。我们可以通过定义绘图r的能量

,  
，

where ρ(vi) is the ith row of R and kρ(vi) − ρ(vj)k2 is the square of the Euclidean length of the line segment joining ρ(vi) and ρ(vj).  
其中，ρ（vi）是r的第i行，kρ（vi）−ρ（vj）k2是连接ρ（vi）和ρ（vj）的线段欧几里得长度的平方。

Then, “good drawings” are drawings that minimize the energy function E. Of course, the trivial representation corresponding to the zero matrix is optimum, so we need to impose extra constraints to rule out the trivial solution.  
那么，“好的图”就是把能量函数e最小化的图，当然，对应于零矩阵的平凡表示是最优的，所以我们需要施加额外的约束来排除平凡解。

We can consider the more general situation where the springs are not necessarily identical. This can be modeled by a symmetric weight (or stiffness) matrix W = (wij), with wij ≥ 0. Then our energy function becomes  
我们可以考虑弹簧不一定相同的更一般的情况。这可以通过对称重量（或刚度）矩阵w=（wij）来建模，wij≥0。然后我们的能量函数变成

.  
.

It turns out that this function can be expressed in terms of the Laplacian L = D − W. The following proposition is shown in Godsil and Royle [77]. We give a slightly more direct proof.  
结果表明，该函数可以用拉普拉斯L=d−w表示。以下命题在Godsil和Royle[77]中给出。我们提供了一个更直接的证据。

Proposition 19.1. Let G = (V,W) be a weighted graph, with |V | = m and W an m × m symmetric matrix, and let R be the matrix of a graph drawing ρ of G in Rn (a m×n matrix).  
提案19.1。设g=（v，w）为加权图，其中v=m，w为m×m对称矩阵，r为Rn中g的图ρ的矩阵（m×n矩阵）。

If L = D − W is the unnormalized Laplacian matrix associated with W, then  
如果l=d−w是与w相关的非正规拉普拉斯矩阵，那么

E(R) = tr(R>LR).  
e（r）=tr（r>lr）。

Proof. Since ρ(vi) is the ith row of R (and ρ(vj) is the jth row of R), if we denote the kth column of R by Rk, using Proposition 18.4, we have  
证据。因为ρ（vi）是r的第i行（而ρ（vj）是r的第j行），如果我们用Rk表示r的第k列，用18.4号命题，我们得到

,  
，

as claimed.   
如要求。

### 19.1. GRAPH DRAWING AND ENERGY MINIMIZATION 19.1。图形绘制与能量最小化

Since the matrix R>LR is symmetric, it has real eigenvalues. Actually, since L is positive semidefinite, so is R>LR. Then the trace of R>LR is equal to the sum of its positive eigenvalues, and this is the energy E(R) of the graph drawing.  
由于矩阵r>lr是对称的，所以它具有实特征值。实际上，因为l是正半定的，所以r>lr也是。那么r>lr的迹线等于它的正特征值之和，这就是图的能量e（r）。

If R is the matrix of a graph drawing in Rn, then for any n×n invertible matrix M, the map that assigns ρ(vi)M to vi is another graph drawing of G, and these two drawings convey the same amount of information. From this point of view, a graph drawing is determined by the column space of R. Therefore, it is reasonable to assume that the columns of R are pairwise orthogonal and that they have unit length. Such a matrix satisfies the equation  
如果r是RN图的矩阵，那么对于任意n×n可逆矩阵m，将ρ（vi）m赋给vi的映射是g的另一个图，这两个图传递的信息量相同。从这个角度来看，图的绘制是由R的列空间决定的，因此，可以合理地假设R的列是成对正交的，并且它们有单位长度。这样的矩阵满足方程

R>R = I.  
r>r=i。

Definition 19.3. If the matrix R of a graph drawing satisfies the equation R>R = I, then the corresponding drawing is called an orthogonal graph drawing.  
定义19.3.如果图的矩阵r满足方程r>r=i，则相应的图称为正交图。

This above condition also rules out trivial drawings. The following result tells us how to find minimum energy orthogonal balanced graph drawings, provided the graph is connected. Recall that  
上述条件也排除了一些琐碎的绘图。下面的结果告诉我们如何找到最小能量正交平衡图，只要图是连接的。回想一下

L1 = 0,  
l1=0，

as we already observed.  
正如我们已经观察到的。

Theorem 19.2. Let G = (V,W) be a weighted graph with |V | = m. If L = D − W is the (unnormalized) Laplacian of G, and if the eigenvalues of L are 0 = λ1 < λ2 ≤ λ3 ≤ ... ≤ λm, then the minimal energy of any balanced orthogonal graph drawing of G in Rn is equal to λ2+···+λn+1 (in particular, this implies that n < m). The m×n matrix R consisting of any unit eigenvectors u2,...,un+1 associated with λ2 ≤ ... ≤ λn+1 yields a balanced orthogonal graph drawing of minimal energy; it satisfies the condition R>R = I.  
定理19.2。设g=（v，w）为v=m的加权图。如果l=d−w是g的（未归一化）拉普拉斯函数，如果l的特征值为0=λ1<λ2≤λ3≤…≤λm，则Rn中g的任意平衡正交图的最小能量等于λ2+·····+λn+1（特别是这意味着n<m）。由任意单位特征向量u2，…，un+1组成的m×n矩阵r与λ2≤…≤λn+1得到最小能量的平衡正交图，满足r>r=i的条件。

Proof. We present the proof given in Godsil and Royle [77] (Section 13.4, Theorem 13.4.1). The key point is that the sum of the n smallest eigenvalues of L is a lower bound for tr(R>LR). This can be shown using a Rayleigh ratio argument; see Proposition 16.25 (the Poincar´e separation theorem). Then any n eigenvectors (u1,...,un) associated with λ1,...,λn achieve this bound. Because the first eigenvalue of L is λ1 = 0 and because we are assuming that λ2 > 0, we have u1 = 1/√m. Since the uj are pairwise orthogonal for i = 2,...,n and since ui is orthogonal to u1 = 1/√m, the entries in ui add up to 0. Consequently, for any ` with 2 ≤ ` ≤ n, by deleting u1 and using (u2,...,u`), we obtain a balanced orthogonal graph drawing in R`−1 with the same energy as the orthogonal graph drawing in R` using (u1,u2,...,u`). Conversely, from any balanced orthogonal drawing in R`−1 using (u2,...,u`), we obtain an orthogonal graph drawing in R` using (u1,u2,...,u`) with the same energy. Therefore, the minimum energy of a balanced orthogonal graph drawing in Rn is equal to the minimum energy of an orthogonal graph drawing in Rn+1, and this minimum is λ2 + ··· + λn+1.   
证据。我们给出了Godsil和Royle[77]中给出的证明（第13.4节，定理13.4.1）。关键是L的n个最小特征值之和是tr（r>lr）的下界。这可以用瑞利比论证来证明；见命题16.25（Poincar'e分离定理）。然后，任何与λ1，…，λn相关的n个特征向量（u1，…，un）都达到这个界限。因为l的第一个特征值是λ1=0，并且因为我们假设λ2>0，所以我们得到了u1=1/√m。由于uj对i=2是正交的，…，n并且因为ui对u1=1/√m是正交的，所以ui中的项加起来是0。因此，对于2≤`≤n的任意‘图，通过删除u1并使用（u2，…，u`），我们得到了r`-1中与r`使用（u1，u2，…，u`）中正交图绘制能量相同的平衡正交图。相反，利用（u2，…，u`）在r`-1中绘制平衡正交图，我们得到了能量相同的r``中使用（u1，u2，…，u`）绘制的正交图。因此，Rn中平衡正交图的最小能量等于Rn+1中正交图的最小能量，此最小能量为λ2+····+λn+1。

Since 1 spans the nullspace of L, using u1 (which belongs to KerL) as one of the vectors in R would have the effect that all points representing vertices of G would have the same first coordinate. This would mean that the drawing lives in a hyperplane in Rn, which is undesirable, especially when n = 2, where all vertices would be collinear. This is why we omit the first eigenvector u1.  
因为1跨越了l的空空间，使用u1（属于kerl）作为r中的向量之一会产生这样的效果，即表示g顶点的所有点都具有相同的第一坐标。这将意味着绘图生活在RN的超平面中，这是不可取的，尤其是当n=2时，所有顶点都将共线。这就是为什么我们省略了第一个特征向量U1。

Observe that for any orthogonal n × n matrix Q, since  
观察任何正交n×n矩阵q，因为

tr(R>LR) = tr(Q>R>LRQ),  
tr（r>lr）=tr（q>r>lrq）

the matrix RQ also yields a minimum orthogonal graph drawing. This amounts to applying the rigid motion Q> to the rows of R.  
矩阵RQ还生成最小正交图。这相当于将刚性运动q>应用于r行。

In summary, if λ2 > 0, an automatic method for drawing a graph in R2 is this:  
总之，如果λ2>0，在r2中绘制图形的自动方法是：

1. Compute the two smallest nonzero eigenvalues λ2 ≤ λ3 of the graph Laplacian L (it is possible that λ3 = λ2 if λ2 is a multiple eigenvalue);  
   计算拉普拉斯L图的两个最小非零特征值λ2≤λ3（如果λ2是多特征值，则可能是λ3=λ2）；
2. Compute two unit eigenvectors u2,u3 associated with λ2 and λ3, and let R = [u2 u3] be the m × 2 matrix having u2 and u3 as columns.  
   计算与λ2和λ3相关的两个单位特征向量u2、u3，并让r=[u2 u3]为m×2矩阵，其中u2和u3为列。
3. Place vertex vi at the point whose coordinates is the ith row of R, that is, (Ri1,Ri2).  
   将顶点vi放在坐标为r的第i行的点上，即（ri1，ri2）。

This method generally gives pleasing results, but beware that there is no guarantee that distinct nodes are assigned distinct images since R can have identical rows. This does not seem to happen often in practice.  
这种方法通常会给出令人满意的结果，但是要注意，由于r可以有相同的行，所以不能保证为不同的节点分配不同的图像。这在实践中似乎并不经常发生。

## 19.2 Examples of Graph Drawings 19.2图表示例

We now give a number of examples using Matlab. Some of these are borrowed or adapted from Spielman [158].  
现在我们用matlab给出一些例子。其中一些是借用或改编自斯皮尔曼[158]。

Example 1. Consider the graph with four nodes whose adjacency matrix is  
例1。考虑具有四个节点的图，其邻接矩阵为

.  
.

We use the following program to compute u2 and u3:  
我们使用以下程序计算u2和u3：

A = [0 1 1 0; 1 0 0 1; 1 0 0 1; 0 1 1 0];  
A=[0 1 1 0；1 0 0 1；1 0 0 1；0 1 1 0]；

D = diag(sum(A));  
d=diag（总和（a））；

L = D - A;  
L=D-A；

[v, e] = eigs(L); gplot(A, v(:,[3 2])) hold on; gplot(A, v(:,[3 2]),’o’)  
[V，E]=EIGS（L）；gplot（A，V（：，[3 2]）保持；gplot（A，V（：，[3 2]），'O'）

### 19.2. EXAMPLES OF GRAPH DRAWINGS 19.2。图形绘图示例

−

0.8

−

0.6

−

0.4

−

0.2

0

0.2

0.4

0.6

0.8

−

0.8

−

0.6

−

0.4

−

0.2

0

0.2

0.4

0.6

0.8

Figure 19.1: Drawing of the graph from Example 1.  
图19.1：示例1中的图形。

The graph of Example 1 is shown in Figure 19.1. The function eigs(L) computes the six largest eigenvalues of L in decreasing order, and corresponding eigenvectors. It turns out that λ2 = λ3 = 2 is a double eigenvalue.  
示例1的图表如图19.1所示。函数特征值（L）按降序计算L的六个最大特征值和相应的特征向量。结果表明，λ2=λ3=2是一个双特征值。

Example 2. Consider the graph G2 shown in Figure 18.3 given by the adjacency matrix  
例2。考虑图18.3所示的图g2，由邻接矩阵给出。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 0 网络错误  1 A = 1 网络错误   网络错误  0 网络错误  0 网络错误 | 1 网络错误  0 网络错误  1 网络错误  1 网络错误  1 网络错误 | 1 网络错误  1 网络错误  0 网络错误  1 网络错误  0 网络错误 | 0 网络错误  1 网络错误  1 网络错误  0 网络错误  1 网络错误 | 0 网络错误  1 网络错误  0. 网络错误  1 0 网络错误 |

We use the following program to compute u2 and u3:  
我们使用以下程序计算u2和u3：

A = [0 1 1 0 0; 1 0 1 1 1; 1 1 0 1 0; 0 1 1 0 1; 0 1 0 1 0];  
A=[0 1 1 0 0；1 0 1 1 1；1 1 0 1 0；0 1 0 1；0 1 0 1 0]；

D = diag(sum(A));  
d=diag（总和（a））；

L = D - A;  
L=D-A；

[v, e] = eig(L); gplot(A, v(:, [2 3])) hold on  
[V，E]=EIG（L）；gplot（A，V（：，[2 3]）保持

gplot(A, v(:, [2 3]),’o’)  
gplot（A，V（：，[2 3]），'O'）

The function eig(L) (with no s at the end) computes the eigenvalues of L in increasing order. The result of drawing the graph is shown in Figure 19.2. Note that node v2 is assigned to the point (0,0), so the difference between this drawing and the drawing in Figure 18.3 is that the drawing of Figure 19.2 is not convex.  
函数eig（l）（结尾没有s）按递增顺序计算l的特征值。绘制图表的结果如图19.2所示。请注意，节点v2被指定为点（0,0），因此此图与图18.3中的图之间的区别在于图19.2的图不是凸的。

Example 3. Consider the ring graph defined by the adjacency matrix A given in the Matlab program shown below:  
例3。考虑以下matlab程序中给出的邻接矩阵a定义的环图：

−

0.8

−

0.6

−

0.4

−

0.2

0

0.2

0.4

0.6

0.8

−

0.8

−

0.6

−

0.4

−

0.2

0

0.2

0.4

0.6

Figure 19.2: Drawing of the graph from Example 2.  
图19.2：示例2中的图表。

A = diag(ones(1, 11),1);  
A=diag（1，11，1）；

A = A + A’;  
A=A+A'；

A(1, 12) = 1; A(12, 1) = 1;  
A（1，12）=1；A（12，1）=1；

D = diag(sum(A));  
d=diag（总和（a））；

L = D - A;  
L=D-A；

[v, e] = eig(L); gplot(A, v(:, [2 3])) hold on  
[V，E]=EIG（L）；gplot（A，V（：，[2 3]）保持

gplot(A, v(:, [2 3]),’o’)  
gplot（A，V（：，[2 3]），'O'）

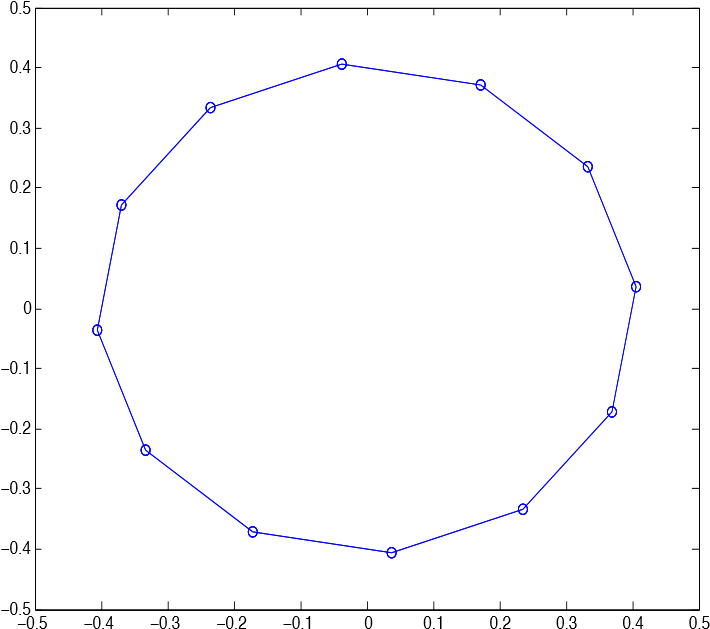


Figure 19.3: Drawing of the graph from Example 3.  
图19.3：示例3中的图形。

Observe that we get a very nice ring; see Figure 19.3. Again λ2 = 0.2679 is a double eigenvalue (and so are the next pairs of eigenvalues, except the last, λ12 = 4).  
观察我们得到了一个非常漂亮的戒指；见图19.3。同样，λ2=0.2679是一个双特征值（下一对特征值也是如此，除了最后一对，λ12=4）。

### 19.2. EXAMPLES OF GRAPH DRAWINGS 19.2。图形绘图示例

Example 4. In this example adapted from Spielman, we generate 20 randomly chosen points in the unit square, compute their Delaunay triangulation, then the adjacency matrix of the corresponding graph, and finally draw the graph using the second and third eigenvalues of the Laplacian.  
例4。在这个由Spielman改编的例子中，我们在单位平方中生成20个随机选择的点，计算它们的Delaunay三角剖分，然后计算相应图的邻接矩阵，最后利用拉普拉斯的第二和第三特征值绘制图。

A = zeros(20,20); xy = rand(20, 2); trigs = delaunay(xy(:,1), xy(:,2)); elemtrig = ones(3) - eye(3); for i = 1:length(trigs),  
A=零（20,20）；xy=rand（20,2）；trigs=delaunay（xy（：，1），xy（：，2））；elemtrig=一（3）-眼（3）；对于i=1：长度（trigs），

A(trigs(i,:),trigs(i,:)) = elemtrig; end  
a（trigs（i，：），trigs（i，：）=elemtrig；结束

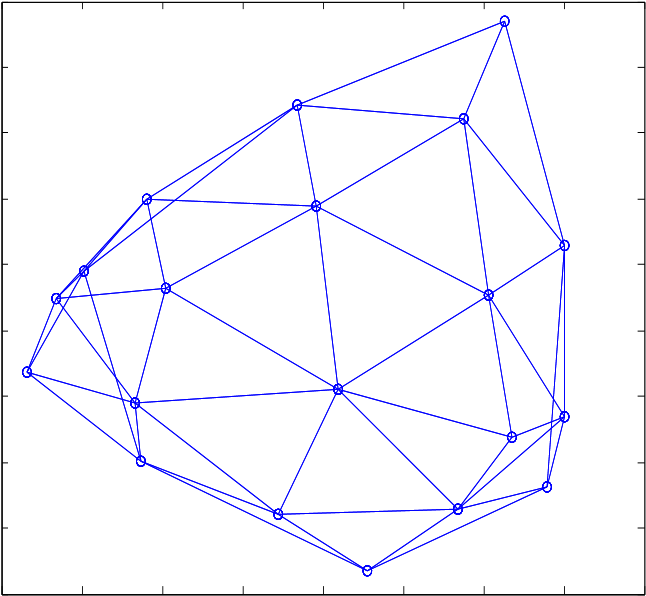
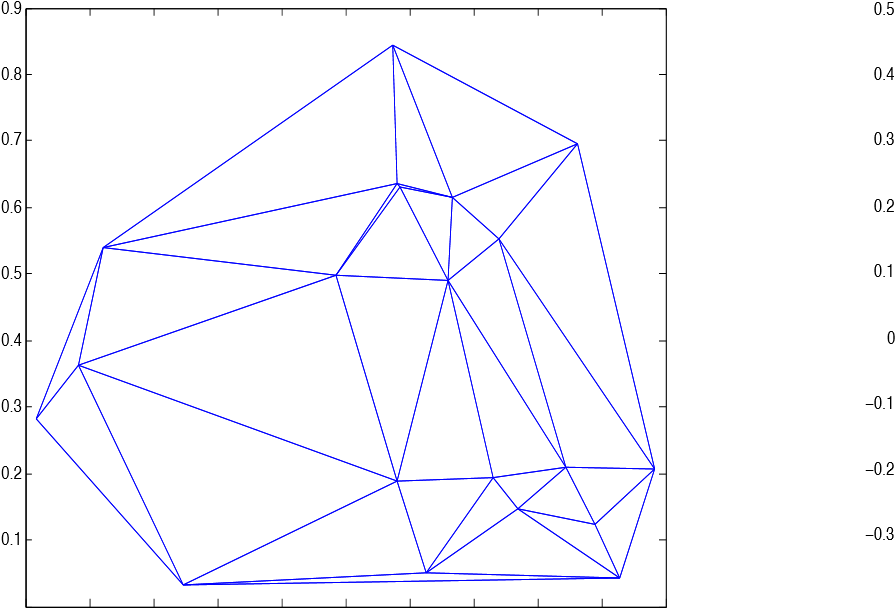
A = double(A >0); gplot(A,xy) D = diag(sum(A));  
A=双（A>0）；gplot（A，xy）D=diag（sum（A））；

L = D - A;  
L=D-A；

[v, e] = eigs(L, 3, ’sm’); figure(2) gplot(A, v(:, [2 1])) hold on  
[V，E]=EIGS（L，3，'SM'）；图（2）gplot（A，V（：，[2 1]）保持

gplot(A, v(:, [2 1]),’o’)  
gplot（A，V（：，[2 1]），'O'）

The Delaunay triangulation of the set of 20 points and the drawing of the corresponding graph are shown in Figure 19.4. The graph drawing on the right looks nicer than the graph on the left but is is no longer planar.  
图19.4显示了20个点集的Delaunay三角测量和相应图形的绘制。右边的图形看起来比左边的图形好，但不再是平面图形。



00 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1 −0.4−0.4 −0.3 −0.2 −0.1 0 0.1 0.2 0.3 0.4  
00 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1−0.4−0.4−0.3−0.2−0.1 0.1 0.2 0.3 0.4

Figure 19.4: Delaunay triangulation (left) and drawing of the graph from Example 4 (right).  
图19.4：Delaunay三角测量（左）和示例4（右）中图形的绘制。

Example 5. Our last example, also borrowed from Spielman [158], corresponds to the skeleton of the “Buckyball,” a geodesic dome invented by the architect Richard Buckminster Fuller (1895–1983). The Montr´eal Biosph`ere is an example of a geodesic dome designed by Buckminster Fuller.  
例5。我们的最后一个例子，也借用了斯皮尔曼（Spielman）的著作[158]，与建筑师理查德·巴克明斯特·富勒（Richard Buckminster Fuller，1895-1983）发明的测地穹顶“Buckyball”的骨架相对应。蒙特利尔生物圈是巴克明斯特富勒设计的测地穹顶的一个例子。

A = full(bucky);  
A=满（Bucky）；

D = diag(sum(A));  
d=diag（总和（a））；

L = D - A;  
L=D-A；

[v, e] = eig(L); gplot(A, v(:, [2 3])) hold on;  
[V，E]=EIG（L）；gplot（A，V（：，[2 3]）保持；

gplot(A,v(:, [2 3]), ’o’)  
gplot（A，V（：，[2 3]），‘O’）

Figure 19.5 shows a graph drawing of the Buckyball. This picture seems a bit squashed for two reasons. First, it is really a 3-dimensional graph; second, λ2 = 0.2434 is a triple eigenvalue. (Actually, the Laplacian of L has many multiple eigenvalues.) What we should really do is to plot this graph in R3 using three orthonormal eigenvectors associated with λ2.  
图19.5显示了Buckyball的图形。这张照片似乎有点压扁，有两个原因。首先，它实际上是一个三维图；其次，λ2=0.2434是一个三重特征值。（实际上，L的拉普拉斯有许多多重特征值。）我们真正应该做的是用三个与λ2相关的正交特征向量在r3中绘制这个图。

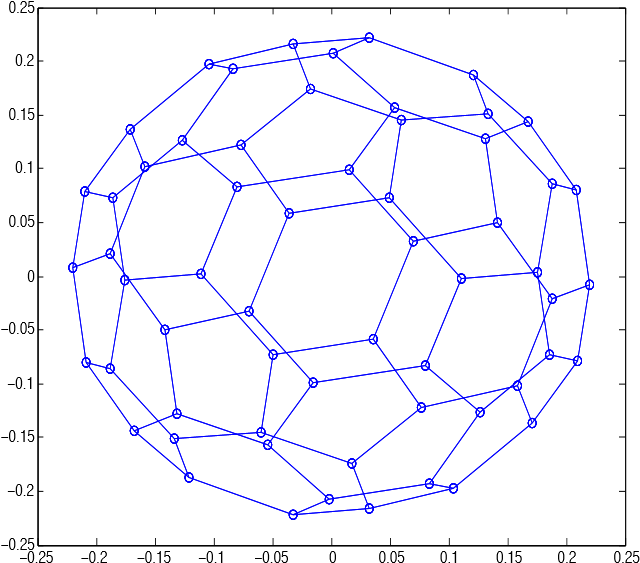


Figure 19.5: Drawing of the graph of the Buckyball.  
图19.5：七叶树图。

A 3D picture of the graph of the Buckyball is produced by the following Matlab program, and its image is shown in Figure 19.6. It looks better!  
下面的matlab程序生成了Buckyball图形的三维图像，其图像如图19.6所示。看起来好多了！

[x, y] = gplot(A, v(:, [2 3]));  
[X，Y]=gplot（A，V（：，[2 3]）；

[x, z] = gplot(A, v(:, [2 4])); plot3(x,y,z)  
[X，Z]=gplot（A，V（：，[2 4]）；plot3（X，Y，Z）

## 19.3 Summary 19.3总结

The main concepts and results of this chapter are listed below:  
本章的主要概念和结果如下：

• Graph drawing.  
•图形绘制。

### 19.3. SUMMARY 19.3。总结

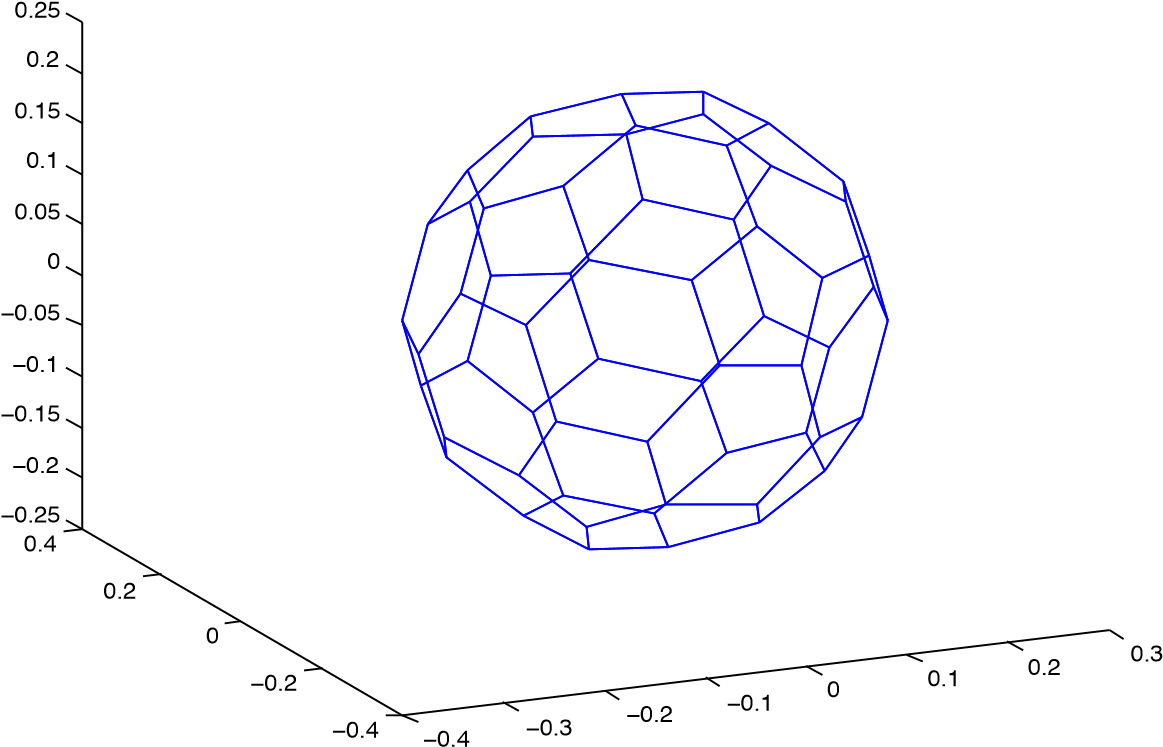


Figure 19.6: Drawing of the graph of the Buckyball in R3.  
图19.6:R3中Buckyball的图形。

* Matrix of a graph drawing.  
  图的矩阵。
* Balanced graph drawing.  
  平衡图绘制。
* Energy E(R) of a graph drawing.  
  图的能量e（r）。
* Orthogonal graph drawing.  
  正交图绘制。
* Delaunay triangulation.  
  Delaunay三角测量。
* Buckyball.  
  巴基鲍尔。

612 CHAPTER 19. SPECTRAL GRAPH DRAWING  
  
612第19章。光谱图绘制

Chapter 20  
第二十章

# Singular Value Decomposition and Polar Form 奇异值分解与极性形式

## 20.1 Properties of f∗ ◦f 20.1 f f的性质

In this section we assume that we are dealing with real Euclidean spaces. Let f : E → E be any linear map. In general, it may not be possible to diagonalize f. We show that every linear map can be diagonalized if we are willing to use two orthonormal bases. This is the celebrated singular value decomposition (SVD). A close cousin of the SVD is the polar form of a linear map, which shows how a linear map can be decomposed into its purely rotational component (perhaps with a flip) and its purely stretching part.  
在本节中，我们假设我们处理的是真正的欧几里得空间。设f:e→e为任意线性映射。一般来说，F不可能对角化。我们证明了如果我们愿意使用两个正交基，每个线性映射都可以对角化。这是著名的奇异值分解（SVD）。SVD的近亲是线性映射的极性形式，它显示了如何将线性映射分解为其纯旋转组件（可能带有翻转）和纯拉伸部分。

The key observation is that f∗ ◦ f is self-adjoint since  
关键的观察是f f自

h(f∗ ◦ f)(u),vi = hf(u),f(v)i = hu,(f∗ ◦ f)(v)i.  
h（f f）（u），v i=hf（u），f（v）i=hu，（f f）（v）i.

Similarly, f ◦ f∗ is self-adjoint.  
同样，f f是自伴的。

The fact that f∗ ◦ f and f ◦ f∗ are self-adjoint is very important, because by Theorem  
F F和F F是自伴的事实非常重要，因为根据定理

16.8In fact, these, it implies thateigenvalues are all nonnegativef∗ ◦f and f ◦f∗ can be diagonalized and that they have real eigenvalues.as shown in the following proposition.  
16.8事实上，这意味着特征值都是非负的f f和f f可以对角化，并且它们具有真正的特征值。如以下命题所示。

Proposition 20.1. The eigenvalues of f∗ ◦ f and f ◦ f∗ are nonnegative. Proof. If u is an eigenvector of f∗ ◦ f for the eigenvalue λ, then  
提案20.1.f f和f f的特征值为非负。证据。如果u是特征值λ的f f的特征向量，那么

h(f∗ ◦ f)(u),ui = hf(u),f(u)i  
h（f\_f）（u），u i=hf（u），f（u）i

and h(f∗ ◦ f)(u),ui = λhu,ui,  
h（f\_f）（u），ui=λhu，ui，

and thus λhu,ui = hf(u),f(u)i,  
因此，λhu，u i=hf（u），f（u）i，

which implies that λ ≥ 0, since h−,−i is positive definite. A similar proof applies to f ◦ f∗.   
这意味着λ≥0，因为h−，−i是正定的。类似的证据也适用于F F。

613  
六百一十三

Thus, the eigenvalues of f∗ ◦f are of the formor 0, where σi > 0, and similarly for f ◦ f∗.  
因此，f f的特征值的形式为0，其中，σi>0，与f f类似。

The above considerations also apply to any linear map f : E → F between two Euclidean spaces (E,h−,−i1) and (F,h−,−i2). Recall that the adjoint f∗ : F → E of f is the unique linear map f∗ such that  
上述考虑也适用于两个欧几里得空间（e，h−、−i1）和（f，h−、−i2）之间的任何线性映射f:e→f。回想一下，f的伴随f：f→e是唯一的线性映射f，这样

hf(u),vi2 = hu,f∗(v)i1, for all u ∈ E and all v ∈ F.  
hf（u），vi2=hu，f（v）i1，对于所有u∈e和所有v∈f。

Then f∗ ◦ f and f ◦ f∗ are self-adjoint (the proof is the same as in the previous case), and the eigenvalues of f∗ ◦ f and f ◦ f∗ are nonnegative.  
则f f和f f为自伴（证明与前一种情况相同），f f和f f的特征值为非负。

Proof. If λ is an eigenvalue of f∗ ◦ f and u (6= 0) is a corresponding eigenvector, we have  
证据。如果λ是f f的特征值，而u（6=0）是对应的特征向量，我们得到

h(f∗ ◦ f)(u),ui1 = hf(u),f(u)i2,  
h（f\_f）（u），ui1=hf（u），f（u）i2，

and also h(f∗ ◦ f)(u),ui1 = λhu,ui1,  
以及h（f\_f）（u），ui1=λhu，ui1，

so  
所以

λhu,ui1,= hf(u),f(u)i2,  
λhu，ui1，=hf（u），f（u）i2，

which implies that λ ≥ 0. A similar proof applies to f ◦ f∗.   
这意味着λ≥0。类似的证据也适用于F F。

The situation is even better, since we will show shortly that f∗ ◦ f and f ◦ f∗ have the same nonzero eigenvalues.  
这种情况甚至更好，因为我们很快就会发现f f和f f具有相同的非零特征值。

Remark: Given any two linear maps f : E → F and g: F → E, where dim(E) = n and dim(F) = m, it can be shown that  
注：任意两个线性映射f:e→f和g:f→e，其中dim（e）=n和dim（f）=m

λm det(λIn − g ◦ f) = λn det(λIm − f ◦ g),  
λm det（λin−g f）=λn det（λim−f g），

and thus g ◦ f and f ◦ g always have the same nonzero eigenvalues; see Problem 14.14.  
因此，g f和f g总是具有相同的非零特征值；参见问题14.14。

eigenvalues ofDefinition 20.1.f∗ ◦Given any linear mapf (and f ◦ f∗) are called thef : E →singular values ofF, the square rootsf. σi > 0 of the positive  
定义的特征值20.1.f给定任何线性mapf（和f f）被取消：e→奇异值关闭，平方根sf。σi>0为正

Definition 20.2. A self-adjoint linear map f : E →is also invertible,E whose eigenvalues are nonnegative isf is said to be positive called positive semidefinite (or positive), and if f definite. In the latter case, every eigenvalue of f is strictly positive.  
定义20.2.自伴线性映射f:e→也是可逆的，其特征值为非负的e称为正半定（或正），如果f定。在后一种情况下，f的每个特征值都是严格正的。

If f : E → F is any linear map, we just showed that f∗ ◦ f and f ◦ f∗ are positive semidefinite self-adjoint linear maps. This fact has the remarkable consequence that every linear map has two important decompositions:  
如果f:e→f是任意线性映射，我们只证明f f和f f是半正定自伴线性映射。这一事实的显著结果是，每个线性映射都有两个重要的分解：

1. The polar form.  
1。极地形态。

### 20.1. PROPERTIES OF f∗ ◦ f 20.1。f f的性质

2. The singular value decomposition (SVD).  
2。奇异值分解（SVD）。

The wonderful thing about the singular value decomposition is that there exist two orthonormal bases (u1,...,un) and (v1,...,vm) such that, with respect to these bases, f is a diagonal matrix consisting of the singular values of f or 0. Thus, in some sense, f can always be diagonalized with respect to two orthonormal bases. The SVD is also a useful tool for solving overdetermined linear systems in the least squares sense and for data analysis, as we show later on.  
奇异值分解的美妙之处在于，存在两个正交基（u1，…，un）和（v1，…，vm），因此，对于这些基，f是由奇异值f或0组成的对角矩阵。因此，在某种意义上，F总是可以相对于两个正交基对角化。SVD也是解决最小二乘意义上的超定线性系统和数据分析的一个有用工具，如我们稍后所示。

First we show some useful relationships between the kernels and the images of f, f∗, f∗ ◦ f, and f ◦ f∗. Recall that if f : E → F is a linear map, the image Imf of f is the subspace f(E) of F, and the rank of f is the dimension dim(Imf) of its image. Also recall that (Theorem 5.11) dim(Kerf) + dim(Imf) = dim(E),  
首先，我们展示了内核与f、f\_、f\_和f\_f\_图像之间的一些有用关系。回想一下，如果f:e→f是一个线性映射，f的图像imf是f的子空间f（e），f的秩是其图像的维数dim（imf）。还记得（定理5.11）dim（切口）+dim（imf）=dim（e），

and that (Propositions 11.11 and 13.13) for every subspace W of E,  
对于e的每个子空间w（命题11.11和13.13），则

dim(W) + dim(W ⊥) = dim(E).  
dim（w）+dim（w）=dim（e）。

Proposition 20.2. Given any two Euclidean spaces E and F, where E has dimension n and F has dimension m, for any linear map f : E → F, we have  
提案20.2.给定任意两个欧几里得空间e和f，其中e有维数n，f有维数m，对于任何线性映射f:e→f，我们有

Kerf = Ker(f∗ ◦ f),  
切口=KER（F F）

Kerf∗ = Ker(f ◦ f∗),  
切口=KER（F F）

Kerf = (Imf∗)⊥,  
切口=（imf），

Kerf∗ = (Imf)⊥, dim(Imf) = dim(Imf∗),  
kerf=（imf），dim（imf）=dim（imf），

and f, f∗, f∗ ◦ f, and f ◦ f∗ have the same rank.  
F、F、F F和F F具有相同的等级。

Proof. To simplify the notation, we will denote the inner products on E and F by the same symbol h−,−i (to avoid subscripts). If f(u) = 0, then (f∗ ◦ f)(u) = f∗(f(u)) = f∗(0) = 0, and so Kerf ⊆ Ker(f∗ ◦ f). By definition of f∗, we have  
证据。为了简化表示法，我们将用相同的符号H−、−I（避免下标）来表示e和f上的内积。如果f（u）=0，那么（f（f）（u）=f（f（u））=f（0）=0，那么切口ker（f f）。根据F的定义，我们有

hf(u),f(u)i = h(f∗ ◦ f)(u),ui  
h f（u），f（u）i=h（f f）（u），用户界面

for all u ∈ E. If (f∗ ◦ f)(u) = 0, since h−,−i is positive definite, we must have f(u) = 0, and so Ker(f∗ ◦ f) ⊆ Kerf. Therefore,  
对于所有的u∈e，如果（f f）（u）=0，因为h−−i是正定的，我们必须有f（u）=0，所以ker（f f）kerf。因此，

Kerf = Ker(f∗ ◦ f).  
切口=KER（F F）。

The proof that Kerf∗ = Ker(f ◦ f∗) is similar.  
切口=KER（F F）的证据相似。

By definition of f∗, we have  
根据F的定义，我们有

hf(u),vi = hu,f∗(v)i for all u ∈ E and all v ∈ F. (∗)  
hf（u），v i=hu，f（v）i表示所有u e和所有v f.（）

This immediately implies that  
这立刻意味着

Kerf = (Imf∗)⊥ and Kerf∗ = (Imf)⊥.  
切口=（imf）和切口=（imf）。

Let us explain why Kerf = (Imf∗)⊥, the proof of the other equation being similar. Because the inner product is positive definite, for every u ∈ E, we have  
让我们解释一下为什么kerf=（imf），这是另一个等式相似的证明。因为内积是正定的，对于每一个u∈e，我们有

* u ∈ Kerf  
  u∈切口
* iff f(u) = 0  
  iff f（u）=0
* iff hf(u),vi = 0 for all v, • by (∗) iff hu,f∗(v)i = 0 for all v,  
  i f f hf（u），v i=0代表所有v，•by（）iff hu，f（v）i=0代表所有v，
* iff u ∈ (Imf∗)⊥.  
  iff u∈（imf）。

Since  
自从

dim(Imf) = n − dim(Kerf)  
dim（imf）=n−dim（切口）

and dim(Imf∗) = n − dim((Imf∗)⊥),  
和dim（imf）=n（imf），

from  
从

Kerf = (Imf∗)⊥  
切口=（imf）

we also have dim(Kerf) = dim((Imf∗)⊥),  
我们还有dim（kerf）=dim（imf），

from which we obtain  
我们从中获得

dim(Imf) = dim(Imf∗).  
dim（imf）=dim（imf）。

Since dim(Ker(f∗ ◦ f)) + dim(Im(f∗ ◦ f)) = dim(E),  
因为dim（ker（f f））+dim（im（f f））=dim（e），

Ker(f∗ ◦ f) = Kerf and Kerf = (Imf∗)⊥, we get  
ker（f f）=kerf和kerf=（imf），我们得到

dim((Imf∗)⊥) + dim(Im(f∗ ◦ f)) = dim(E).  
dim（（im f））+dim（im（f f））=dim（e）。

Since dim((Imf∗)⊥) + dim(Imf∗) = dim(E),  
因为dim（（imf））+dim（imf）=dim（e），

we deduce that dim(Imf∗) = dim(Im(f∗ ◦ f)).  
我们推导出dim（im f）=dim（im（f f））。

A similar proof shows that  
类似的证据表明

dim(Imf) = dim(Im(f ◦ f∗)).  
dim（im f）=dim（im（f\_f）。

Consequently, f, f∗, f∗ ◦ f, and f ◦ f∗ have the same rank.   
因此，F、F、F F和F F具有相同的等级。

20.2. SINGULAR VALUE DECOMPOSITION FOR SQUARE MATRICES  
20.2。方阵的奇异值分解

## 20.2 Singular Value Decomposition for Square Matrices 20.2平方矩阵的奇异值分解

We will now prove that every square matrix has an SVD. Stronger results can be obtained if we first consider the polar form and then derive the SVD from it (there are uniqueness properties of the polar decomposition). For our purposes, uniqueness results are not as important so we content ourselves with existence results, whose proofs are simpler. Readers interested in a more general treatment are referred to Gallier [73].  
现在我们将证明每个平方矩阵都有一个SVD。如果我们先考虑极性形式，然后从中导出SVD（极性分解具有唯一性），则可以得到更强有力的结果。就我们的目的而言，唯一性结果并没有那么重要，所以我们满足于存在结果，其证明更简单。读者对更普遍的治疗感兴趣，请参考加利尔[73]。

The early history of the singular value decomposition is described in a fascinating paper by Stewart [160]. The SVD is due to Beltrami and Camille Jordan independently (1873, 1874). Gauss is the grandfather of all this, for his work on least squares (1809, 1823) (but Legendre also published a paper on least squares!). Then come Sylvester, Schmidt, and Hermann Weyl. Sylvester’s work was apparently “opaque.” He gave a computational method to find an SVD. Schmidt’s work really has to do with integral equations and symmetric and asymmetric kernels (1907). Weyl’s work has to do with perturbation theory (1912). Autonne came up with the polar decomposition (1902, 1915). Eckart and Young extended SVD to rectangular matrices (1936, 1939).  
史都华[160]在一篇引人入胜的论文中描述了奇异值分解的早期历史。SVD独立于贝尔特拉米和卡米尔·乔丹（18731874）。高斯是所有这些的祖父，因为他在最小二乘（18091823）上的工作（但勒让德也发表了一篇关于最小二乘的论文！）然后是西尔维斯特、施密特和赫尔曼·韦尔。西尔维斯特的工作显然是“不透明的”，他给出了一种计算方法来寻找SVD。施密特的工作实际上与积分方程和对称和不对称内核有关（1907年）。韦尔的工作与微扰理论（1912）有关。Autonne提出了极地分解（19021915）。Eckart和Young将SVD扩展到了矩形矩阵（1936、1939）。

Theorem 20.3. (Singular value decomposition) For every real n×n matrix A there are two orthogonal matrices U and V and a diagonal matrix D such that A = V DU>, where D is of  
定理20.3。（奇异值分解）对于每个实n×n矩阵a，有两个正交矩阵u和v和一个对角矩阵d，其中a=v du>，其中d是

|  |  |  |  |
| --- | --- | --- | --- |
| the form 网络错误 |  |  |  |
|  网络错误  σ1 网络错误   网络错误  D =  ... 网络错误   网络错误   网络错误 | σ2 网络错误  ... 网络错误 | ... 网络错误  ... ... ... 网络错误 |  网络错误   网络错误  ... , 网络错误  σn 网络错误 |

where σ1,...,σr are the singular values of f, i.e., the (positive) square roots of the nonzero eigenvalues of A>A and AA>, and σr+1 = ··· = σn = 0. The columns of U are eigenvectors of A>A, and the columns of V are eigenvectors of AA>.  
式中，σ1，…，σr是f的奇异值，即a>a和aa>的非零特征值的（正）平方根，以及σr+1=·····=σn=0。u列是a>a的特征向量，v列是aa>的特征向量。

Proof. Since A>A is a symmetric matrix, in fact, a positive semidefinite matrix, there exists an orthogonal matrix U such that  
证据。由于a>a是对称矩阵，实际上是半正定矩阵，因此存在一个正交矩阵u，使得

A>A = UD2U>,  
a>a=ud2u>，

with D = diag(σ1,...,σr,0,...,0), where are the nonzero eigenvalues of A>A, and where r is the rank of A; that is, σ1,...,σr are the singular values of A. It follows that  
d=diag（σ1，…，σr，0，…，0），其中a>a的非零特征值，其中r是a的秩；即，σ1，…，σr是a的奇异值，其结果如下：

U>A>AU = (AU)>AU = D2,  
u>a>au=（au）>au=d2，

and if we let fj be the jth column of AU for j = 1,...,n, then we have  
如果我们让fj是j=1，…，n的au的jth列，那么我们有

hfi,fji = σi2δij, 1 ≤ i,j ≤ r  
hfi，fji=σi2δi j，1≤i，j≤r

and  
和

fj = 0, r + 1 ≤ j ≤ n.  
fj=0，r+1≤j≤n。

If we define (v1,...,vr) by  
如果我们定义（v1，…，vr）

vj = σj−1fj, 1 ≤ j ≤ r,  
Vj=σj−1fj，1≤j≤r，

then we have  
然后我们有了

hvi,vji = δij, 1 ≤ i,j ≤ r,  
hvi，vji=δi j，1≤i，j≤r，

so complete (v1,...,vr) into an orthonormal basis (v1,...,vr,vr+1,...,vn) (for example, using Gram–Schmidt). Now since fj = σjvj for j = 1...,r, we have  
因此，将（v1，…，vr）完全转换为正态基（v1，…，vr，vr+1，…，vn）（例如，使用gram-schmidt）。既然fj=σjvj，对于j=1…，r，我们有

hvi,fji = σjhvi,vji = σjδi,j, 1 ≤ i ≤ n, 1 ≤ j ≤ r  
hvi，fji=σj hvi，vji=σjδi，j，1≤i≤n，1≤j≤r

and since fj = 0 for j = r + 1,...,n,  
既然j=r+1时fj=0，…，n，

hvi,fji = 0 1 ≤ i ≤ n, r + 1 ≤ j ≤ n.  
hvi，fji=0 1≤i≤n，r+1≤j≤n。

If V is the matrix whose columns are v1,...,vn, then V is orthogonal and the above equations prove that  
如果v是列为v1，…，vn的矩阵，则v是正交的，上述方程证明

V >AU = D,  
v>au=d，

which yields A = V DU>, as required.  
根据需要，得出a=v du>。

The equation A = V DU> implies that  
方程式a=v du>意味着

A>A = UD2U>, AA> = V D2V >,  
a>a=ud2u>，aa>=v d2v>，

which shows that A>A and AA> have the same eigenvalues, that the columns of U are eigenvectors of A>A, and that the columns of V are eigenvectors of AA>.   
这表明A>A和AA>具有相同的特征值，U列是A>A的特征向量，V列是AA>的特征向量。

Example 20.1. Here is a simple example of how to use the proof of Theorem 20.3 to obtain an SVD decomposition. Let . Then , and  
例20.1。下面是一个简单的例子，说明如何使用定理20.3的证明来获得SVD分解。让。然后，和

. A simple calculation shows that the eigenvalues of A>A are 2 and 0, and  
. 简单计算表明，a>a的特征值为2和0，以及

for the eigenvalue 2, a unit eigenvector is , while a unit eigenvector for the eigenvalue  
对于特征值2，单位特征向量为，而单位特征向量为

. Observe that the singular values are σ1 = √2 and σ2 = 0. Furthermore,  
. 观察奇异值为σ1=√2和σ2=0。此外，

. To determine V , the proof of Theorem 20.3 tells us to first  
. 为了确定v，定理20.3的证明告诉我们首先

calculate  
计算

and then set  
然后设置

.  
.

### 20.2. SINGULAR VALUE DECOMPOSITION FOR SQUARE MATRICES 20.2。方阵的奇异值分解

Once v1 is determined, since σ2 = 0, we have the freedom to choose v2 such that (v1,v2) forms an orthonormal basis for R2. Naturally, we chose and set. Of course we could have found V by directly computing the eigenvalues and eigenvectors for AA>. We leave it to the reader to check that  
一旦确定了v1，因为σ2=0，我们就可以自由地选择v2，这样（v1，v2）就形成了r2的正交基。当然，我们选择和设置。当然，我们可以通过直接计算aa>的特征值和特征向量找到v。我们把它留给读者检查一下

.  
.

Theorem 20.3 suggests the following definition.  
定理20.3给出了以下定义。

Definition 20.3. A triple (U,D,V ) such that A = V D U>, where U and V are orthogonal and D is a diagonal matrix whose entries are nonnegative (it is positive semidefinite) is called a singular value decomposition (SVD) of A.  
定义20.3.一种三重矩阵（u，d，v），其中a=v d u>，其中u和v是正交的，d是一个项为非负（正半定）的对角矩阵，称为a的奇异值分解（svd）。

The Matlab command for computing an SVD A = V DU> of a matrix A is [V, D, U] = svd(A).  
用于计算矩阵A的svd a=v d u>的matlab命令是[v，d，u]=svd（a）。

The proof of Theorem 20.3 shows that there are two orthonormal bases (u1,...,un) and (v1,...,vn), where (u1,...,un) are eigenvectors of A>A and (v1,...,vn) are eigenvectors of AA>. Furthermore, (u1,...,ur) is an orthonormal basis of ImA>, (ur+1,...,un) is an orthonormal basis of KerA, (v1,...,vr) is an orthonormal basis of ImA, and (vr+1,...,vn) is an orthonormal basis of KerA>.  
定理20.3的证明表明，有两个正交基（u1，…，un）和（v1，…，vn），其中（u1，…，un）是a>a的特征向量，（v1，…，vn）是aa>的特征向量。此外，（u1，…，ur）是ima>的正态基础，（ur+1，…，un）是kera的正态基础，（v1，…，vr）是ima的正态基础，（vr+1，…，vn）是kera>的正态基础。

Using a remark made in Chapter 4, if we denote the columns of U by u1,...,un and the columns of V by v1,...,vn, then we can write  
用第四章的注释，如果我们用u1，…，un表示u的列，用v1，…，vn表示v的列，那么我们可以写

.  
.

As a consequence, if r is a lot smaller than n (we write ), we see that A can be reconstructed from U and V using a much smaller number of elements. This idea will be used to provide “low-rank” approximations of a matrix. The idea is to keep only the k top singular values for some suitable for which σk+1,...σr are very small.  
因此，如果r比n小很多（我们写的），我们可以看到a可以用更少的元素从u和v重建。这个想法将被用来提供矩阵的“低阶”近似值。我们的想法是，对于一些σk+1，…σr非常小的情况，只保留k的顶部奇异值。

Remarks:  
评论：

1. In Strang [165] the matrices U,V,D are denoted by U = Q2, V = Q1, and D = Σ, and an SVD is written as. This has the advantage that Q1 comes before Q2 in  
   在Strang[165]中，矩阵u、v、d用u=q2、v=q1和d=∑表示，SVD写为。这具有q1在q2之前的优势。

. This has the disadvantage that A maps the columns of Q2 (eigenvectors  
. 这有一个缺点，即A映射了q2列（特征向量

of A>A) to multiples of the columns of Q1 (eigenvectors of AA>).  
a>a）到q1列的倍数（aa>的特征向量）。

1. Algorithms for actually computing the SVD of a matrix are presented in Golub and Van Loan [80], Demmel [49], and Trefethen and Bau [171], where the SVD and its applications are also discussed quite extensively.  
   实际计算矩阵SVD的算法在Golub和van Loan[80]、Demmel[49]和Trefethen和Bau[171]中给出，其中，SVD及其应用也得到了广泛讨论。
2. If A is a symmetric matrix, then in general, there is no SVD V ΣU> of A with V = U. However, if A is positive semidefinite, then the eigenvalues of A are nonnegative, and so the nonzero eigenvalues of A are equal to the singular values of A and SVDs of A are of the form  
   如果A是一个对称矩阵，那么一般来说，在V=U的情况下，A的SVD V∑U>是不存在的，但是，如果A是半正定的，那么A的特征值是非负的，因此A的非零特征值等于A的奇异值，A的SVD的形式是
   1. = V ΣV >.  
      =V∑V>。
3. The SVD also applies to complex matrices. In this case, for every complex n×n matrix  
   SVD也适用于复杂矩阵。在这种情况下，对于每个复杂的n×n矩阵

A, there are two unitary matrices U and V and a diagonal matrix D such that  
有两个单位矩阵u和v和一个对角矩阵d，这样

* 1. = V D U∗,  
     =v d u，

where D is a diagonal matrix consisting of real entries σ1,...,σn, where σ1,...,σr are the singular values of A, i.e., the positive square roots of the nonzero eigenvalues of A∗A and AA∗, and σr+1 = ... = σn = 0.  
式中，d是由实数项σ1，…，σn组成的对角矩阵，其中，σ1，…，σr是a的奇异值，即a a和aa的非零特征值的正平方根，以及σr+1=。=σn=0。

## 20.3 Polar Form for Square Matrices 20.3方阵的极性形式

A notion closely related to the SVD is the polar form of a matrix.  
与SVD密切相关的概念是矩阵的极性形式。

Definition 20.4. A pair (R,S) such that A = RS with R orthogonal and S symmetric positive semidefinite is called a polar decomposition of A.  
定义20.4.具有r正交和s对称半正定的a=rs的对（r，s）称为a的极分解。

Theorem 20.3 implies that for every real n×n matrix A, there is some orthogonal matrix  
定理20.3表明，对于每个实n×n矩阵a，都有一些正交矩阵

R and some positive semidefinite symmetric matrix S such that  
r和一些半正定对称矩阵s，这样

A = RS.  
A=卢比。

This is easy to show and we will prove it below. Furthermore, R,S are unique if A is invertible, but this is harder to prove; see Problem 20.9.  
这很容易展示，我们将在下面证明。此外，如果a是可逆的，r，s是唯一的，但这很难证明；见问题20.9。

For example, the matrix  
例如，矩阵

− −  
−−

is both orthogonal and symmetric, and A = RS with R = A and S = I, which implies that some of the eigenvalues of A are negative.  
是正交和对称的，A=r s，R=a和S=i，这意味着A的一些特征值是负的。

Remark: In the complex case, the polar decomposition states that for every complex n×n matrix A, there is some unitary matrix U and some positive semidefinite Hermitian matrix H such that  
注：在复杂情况下，极分解表明，对于每一个复杂的n×n矩阵a，都有一些幺正矩阵u和一些半正定厄米特矩阵h，这样

A = UH.  
A=呃。

### 20.3. POLAR FORM FOR SQUARE MATRICES 20.3。方阵的极性形式

It is easy to go from the polar form to the SVD, and conversely.  
从极性形式到SVD很容易，反之亦然。

Given an SVD decomposition A = V D U>, let R = V U> and S = UD U>. It is clear that R is orthogonal and that S is positive semidefinite symmetric, and  
给定SVD分解a=v d u>，设r=v u>和s=ud u>。很明显，r是正交的，s是半正定对称的，并且

RS = V U>UD U> = V D U> = A.  
rs=v u>u d u>=v d u>=a。

Example 20.2. Recall from Example 20.1 that A = V DU> where V = I2 and  
例20.2。从示例20.1中回忆，a=v du>其中v=i2和

.  
.

Set R = V U> = U and  
设置r=v u>=u和

.  
.

Since has eigenvalues √2 and 0. We leave it to the reader to check that A = RS.  
因为有特征值√2和0。我们把它留给读者来检查a=rs。

Going the other way, given a polar decomposition A = R1S, where R1 is orthogonal and S is positive semidefinite symmetric, there is an orthogonal matrix R2 and a positive semidefinite diagonal matrix D such that, and thus  
另一方面，给定极分解a=r1 s，其中r1是正交的，s是半正定对称的，有一个正交矩阵r2和一个半正定对角矩阵d，因此

,  
，

where V = R1R2 and U = R2 are orthogonal.  
其中v=R1r2和u=r2是正交的。

Example 20.3. Let and A = R1S, where and S =  
例20.3。Let和a=R1s，其中和s=

. This is the polar decomposition of Example 20.2. Observe that  
. 这是实施例20.2的极性分解。注意

.  
.

Set U = R2 and to obtain the SVD decomposition of Example 20.1.  
设置u=r2并获得实施例20.1的SVD分解。

The eigenvalues and the singular values of a matrix are typically not related in any obvious way. For example, the n × n matrix  
矩阵的特征值和奇异值通常没有任何明显的联系。例如，n×n矩阵

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  网络错误  1 网络错误  0 网络错误   网络错误  0 A = ... 网络错误   网络错误   网络错误  0 网络错误   网络错误  0 网络错误   网络错误  0 网络错误 | 2 网络错误  1 网络错误  0 网络错误  ... 网络错误  0 网络错误  0 网络错误  0 网络错误 | 0 网络错误  2 网络错误  1 网络错误  ... 网络错误  ... 网络错误  ... 网络错误  ... 网络错误 | 0 网络错误  0 网络错误  2 网络错误  ... 网络错误  0 网络错误  0 网络错误  0 网络错误 | ... 网络错误  ... 网络错误  ... ... 网络错误  1 网络错误  0 网络错误  0 网络错误 | 0 网络错误  0 网络错误  0 网络错误  ... 网络错误  2 网络错误  1 网络错误  0 网络错误 | 0 网络错误  0 网络错误  0 网络错误  ... 网络错误  0 网络错误  2 1 网络错误 |

has the eigenvalue 1 with multiplicity n, but its singular values, σ1 ≥ ··· ≥ σn, which are the positive square roots of the eigenvalues of the matrix B = A>A with  
具有多重性n的特征值1，但其奇异值，σ1≥········································

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| have a wide spread, since 网络错误 |  网络错误  1 网络错误  2 网络错误   网络错误  0 B = ... 网络错误   网络错误   网络错误  0 网络错误   网络错误  0 网络错误   网络错误  0 网络错误 | 2 网络错误  5 网络错误  2 网络错误  ... 网络错误  0 网络错误  0 网络错误  0 网络错误 | 0 网络错误  2 网络错误  5 网络错误  ... 网络错误  ... 网络错误  ... 网络错误  ... 网络错误 | 0 网络错误  0 网络错误  2 网络错误  ... 网络错误  2 网络错误  0 网络错误  0 网络错误 | ... 网络错误  ... 网络错误  ... ... 网络错误  5 网络错误  2 网络错误  0 网络错误 | 0 网络错误  0 网络错误  0 网络错误  ... 网络错误  2 网络错误  5 网络错误  2 网络错误 | 0 网络错误  0 网络错误  0 网络错误  ... 网络错误  0 网络错误  2 5 网络错误 |

= cond2(A) ≥ 2n−1.  
=cond2（a）≥2n−1。

If A is a complex n × n matrix, the eigenvalues λ1,...,λn and the singular values σ1 ≥ σ2 ≥ ··· ≥ σn of A are not unrelated, since  
如果a是一个复n×n矩阵，a的特征值λ1，…，λn和奇异值σ1≥σ2≥·································



and  
和

|λ1|···|λn| = |det(A)|,  
|λ1···λn=Det（a），

so we have  
所以我们有

|λ1|···|λn| = σ1 ···σn.  
|λ1···λn=σ1····σn.

More generally, Hermann Weyl proved the following remarkable theorem:  
一般来说，赫尔曼·韦尔证明了以下显著定理：

Theorem 20.4. (Weyl’s inequalities, 1949) For any complex n×n matrix, A, if λ1,...,λn ∈ C are the eigenvalues of A and σ1,...,σn ∈ R+ are the singular values of A, listed so that  
定理20.4.（Weyl's不等式，1949）对于任何复杂的n×n矩阵，a，如果λ1，…，λn∈c是a的特征值和σ1，…，σn∈r+是a的奇异值，列出如下：

|λ1| ≥ ··· ≥ |λn| and σ1 ≥ ··· ≥ σn ≥ 0, then  
|λ1≥······≥λn和σ1≥········≥σn≥0，则

|λ1|···|λn| = σ1 ···σn and  
|λ1···λn=σ1·····σn和

|λ1|···|λk| ≤ σ1 ···σk, for k = 1,...,n − 1.  
|λ1···λk≤σ1······σk，对于k=1，…，n−1。

A proof of Theorem 20.4 can be found in Horn and Johnson [93], Chapter 3, Section 3.3, where more inequalities relating the eigenvalues and the singular values of a matrix are given.  
定理20.4的证明可在Horn和Johnson[93]第3章第3.3节中找到，其中给出了更多关于矩阵特征值和奇异值的不等式。

Theorem 20.3 can be easily extended to rectangular m × n matrices, as we show in the next section. For various versions of the SVD for rectangular matrices, see Strang [165] Golub and Van Loan [80], Demmel [49], and Trefethen and Bau [171].  
定理20.3可以很容易地扩展到矩形m×n矩阵，如我们在下一节所示。有关矩形矩阵的SVD的各种版本，请参见Strang[165]Golub和van Loan[80]、Demmel[49]和Trefethen和Bau[171]。

20.4. SINGULAR VALUE DECOMPOSITION FOR RECTANGULAR MATRICES  
20.4。矩形矩阵的奇异值分解

## 20.4 Singular Value Decomposition for Rectangular Matrices 20.4矩形矩阵的奇异值分解

Here is the generalization of Theorem 20.3 to rectangular matrices.  
这是定理20.3对矩形矩阵的推广。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| two orthogonal matrices U (n×n) and V (m×m) and a diagonal m×n matrix D such that 网络错误 | | | | | |
| A = V D U>, where D is of the form 网络错误 |  |  |  |  |  |
| σ1 ...  网络错误   σ2 ...  网络错误   ... ... ... ...  σ1 网络错误   网络错误     网络错误  D =  . ... σn or D =  ... 网络错误   0 .. ... 0   网络错误    网络错误   ... ... ... ...  网络错误   网络错误   0 ... ... 0  网络错误 | σ2 网络错误  ... 网络错误 | ... 网络错误  ... ... ... 网络错误 | ... 网络错误  σm 网络错误 | 0 ... 网络错误  0 ... 网络错误  . 网络错误  0 .. 网络错误  0 ... 网络错误 | 0 网络错误  0 , 网络错误  0 0 网络错误 |

Theorem 20.5. (Singular value decomposition) For every real m × n matrix A, there are  
定理20.5。（奇异值分解）对于每个实M×N矩阵A，有

where σ1,...,σr are the singular values of f, i.e. the (positive) square roots of the nonzero eigenvalues of A>A and AA>, and σr+1 = ... = σp = 0, where p = min(m,n). The columns of U are eigenvectors of A>A, and the columns of V are eigenvectors of AA>.  
其中，σ1，…，σr是f的奇异值，即a>a和aa>的非零特征值的（正）平方根，以及σr+1=。=σp=0，其中p=min（m，n）。u列是a>a的特征向量，v列是aa>的特征向量。

Proof. As in the proof of Theorem 20.3, since A>A is symmetric positive semidefinite, there exists an n × n orthogonal matrix U such that  
证据。在定理20.3的证明中，由于a>a是对称半正定的，所以存在一个n×n正交矩阵u，使得

A>A = UΣ2U>,  
A>A=U∑2U>，

with Σ = diag(σ1,...,σr,0,...,0), where are the nonzero eigenvalues of A>A, and where r is the rank of A. Observe that r ≤ min{m,n}, and AU is an m × n matrix. It follows that  
当∑=diag（σ1，…，σr，0，…，0），其中是a>a的非零特征值，其中r是a的秩。观察r≤min m，n，au是m×n矩阵。接下来是

U>A>AU = (AU)>AU = Σ2,  
u>a>au=（au）>au=∑2，

and if we let fj ∈ Rm be the jth column of AU for j = 1,...,n, then we have  
如果我们让Fj∈Rm为j=1，…，n的au的jth列，那么我们得到

hfi,fji = σi2δij, 1 ≤ i,j ≤ r  
hfi，fji=σi2δi j，1≤i，j≤r

and  
和

fj = 0, r + 1 ≤ j ≤ n.  
fj=0，r+1≤j≤n。

If we define (v1,...,vr) by  
如果我们定义（v1，…，vr）

vj = σj−1fj, 1 ≤ j ≤ r,  
Vj=σj−1fj，1≤j≤r，

then we have  
然后我们有了

hvi,vji = δij, 1 ≤ i,j ≤ r,  
hvi，vji=δi j，1≤i，j≤r，

so complete (v1,...,vr) into an orthonormal basis (v1,...,vr,vr+1,...,vm) (for example, using Gram–Schmidt).  
因此，将（v1，…，vr）完全转换为正态基（v1，…，vr，vr+1，…，vm）（例如，使用gram-schmidt）。

Now since fj = σjvj for j = 1...,r, we have  
既然fj=σjvj，对于j=1…，r，我们有

hvi,fji = σjhvi,vji = σjδi,j, 1 ≤ i ≤ m, 1 ≤ j ≤ r  
hvi，fji=σjhvi，vji=σjδi，j，1≤i≤m，1≤j≤r

and since fj = 0 for j = r + 1,...,n, we have  
既然j=r+1时fj=0，…，n，我们有

hvi,fji = 0 1 ≤ i ≤ m, r + 1 ≤ j ≤ n.  
hvi，fji=0 1≤i≤m，r+1≤j≤n。

If V is the matrix whose columns are v1,...,vm, then V is an m×m orthogonal matrix and if m ≥ n, we let  
如果v是列为v1，…，v m的矩阵，那么v是m×m正交矩阵，如果m≥n，我们将

,  
，

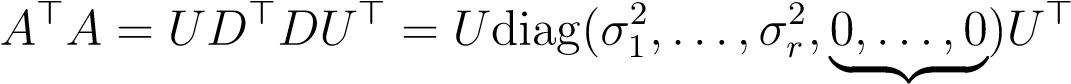
else if n ≥ m, then we let  
否则，如果n≥m，那么我们让

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  网络错误  σ1 网络错误   网络错误  D =  ... 网络错误   网络错误   网络错误 | σ2 网络错误  ... 网络错误 | ... 网络错误  ... ... 网络错误  ... 网络错误 | ... 网络错误  σm 网络错误 | 0 网络错误  0 网络错误  0 网络错误  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 |

In either case, the above equations prove that  
无论哪种情况，上述方程都证明了

V >AU = D,  
v>au=d，

which yields A = V DU>, as required. The equation A = V DU> implies that  
根据需要，得出a=v du>。方程式a=v du>意味着



n−r  
n r

and  
和

AA> = V DD>V > = V diag(,  
aa>=v dd>v>=v诊断（，

m−r  
对氨基苯甲酸

which shows that A>A and AA> have the same nonzero eigenvalues, that the columns of U are eigenvectors of A>A, and that the columns of V are eigenvectors of AA>.   
这表明A>A和AA>具有相同的非零特征值，U列是A>A的特征向量，V列是AA>的特征向量。

A triple (U,D,V ) such that A = V D U> is called a singular value decomposition (SVD) of A.  
一种三重（u，d，v），使a=v，d，u>称为a的奇异值分解（svd）。

### 20.4. SINGULAR VALUE DECOMPOSITION FOR RECTANGULAR MATRICES 20.4。矩形矩阵的奇异值分解

Example 20.4. Let . Then, and AA> =  
例20.4。让。然后，和aa>。=

. The reader should verify that A>A = UΣ2U> where Σ and  
.读者应确认a>a=u∑2u>其中∑和

. Since ,  
. 从那以后，

and complete an orthonormal basis for R3 by assigning, and. Thus  
并通过赋值完成r3的正交基，和。因此

V = I3, and the reader should verify that A = V DU>, where.  
v=i3，读卡器应验证a=v du>，其中。

Even though the matrix D is an m×n rectangular matrix, since its only nonzero entries are on the descending diagonal, we still say that D is a diagonal matrix.  
尽管矩阵d是一个m×n的矩形矩阵，但由于它的唯一非零项是在降对角线上，所以我们仍然认为d是一个对角线矩阵。

The Matlab command for computing an SVD A = V DU> of a matrix A is also [V, D, U] = svd(A).  
用于计算矩阵A的svd a=v d u>的matlab命令也是[v，d，u]=svd（a）。

If we view A as the representation of a linear map f : E → F, where dim(E) = n and dim(F) = m, the proof of Theorem 20.5 shows that there are two orthonormal bases (u1,..., un) and (v1,...,vm) for E and F, respectively, where (u1,...,un) are eigenvectors of f∗ ◦ f and (v1,...,vm) are eigenvectors of f ◦f∗. Furthermore, (u1,...,ur) is an orthonormal basis of Imf∗, (ur+1,...,un) is an orthonormal basis of Kerf, (v1,...,vr) is an orthonormal basis of Imf, and (vr+1,...,vm) is an orthonormal basis of Kerf∗.  
如果我们把a看作线性映射f:e→f的表示，其中dim（e）=n和dim（f）=m，定理20.5的证明表明e和f分别有两个正交基（u1，…，un）和（v1，…，vm），其中（u1，…，un）是f f和（v1，…，vm）的特征向量是特征向量。F F的任务大纲。此外，（u1，…，ur）是imf的正态基，（ur+1，…，un）是kerf的正态基，（v1，…，vr）是imf的正态基，（vr+1，…，vm）是kerf的正态基。

The SVD of matrices can be used to define the pseudo-inverse of a rectangular matrix; we will do so in Chapter 21. The reader may also consult Strang [165], Demmel [49], Trefethen and Bau [171], and Golub and Van Loan [80].  
矩阵的SVD可用于定义矩形矩阵的伪逆矩阵；我们将在第21章中这样做。读者还可以咨询Strang[165]、Demmel[49]、Trefetten和Bau[171]以及Golub和van Loan[80]。

One of the spectral theorems states that a symmetric matrix can be diagonalized by an orthogonal matrix. There are several numerical methods to compute the eigenvalues of a symmetric matrix A. One method consists in tridiagonalizing A, which means that there exists some orthogonal matrix P and some symmetric tridiagonal matrix T such that A = PTP >. In fact, this can be done using Householder transformations; see Theorem 22.2. It is then possible to compute the eigenvalues of T using a bisection method based on Sturm sequences. One can also use Jacobi’s method. For details, see Golub and Van Loan [80], Chapter 8, Demmel [49], Trefethen and Bau [171], Lecture 26, Ciarlet [41], and Chapter 22. Computing the SVD of a matrix A is more involved. Most methods begin by finding orthogonal matrices U and V and a bidiagonal matrix B such that A = V BU>; see Problem 12.8 and Problem 20.3. This can also be done using Householder transformations. Observe that B>B is symmetric tridiagonal. Thus, in principle, the previous method to diagonalize a symmetric tridiagonal matrix can be applied. However, it is unwise to compute B>B explicitly, and more subtle methods are used for this last step; the matrix of Problem 20.1 can be used, and see Problem 20.3. Again, see Golub and Van Loan [80], Chapter 8, Demmel [49], and Trefethen and Bau [171], Lecture 31.  
谱定理之一指出对称矩阵可以由正交矩阵对角化。计算对称矩阵A特征值的数值方法有多种，其中一种方法是三对角化A，即存在一些正交矩阵P和一些对称的三对角矩阵T，从而a=p t p>。事实上，这可以通过户主转换来实现；参见定理22.2。然后可以使用基于sturm序列的二分法计算t的特征值。也可以使用雅可比方法。有关详细信息，请参见Golub和van Loan[80]、第8章、Demmel[49]、Trefethen和Bau[171]、第26讲、Ciarlet[41]和第22章。计算矩阵A的SVD更为复杂。大多数方法首先找到正交矩阵u和v以及双对角矩阵b，使a=v bu>；参见问题12.8和问题20.3。这也可以通过户主转换来实现。观察b>b为对称三对角。因此，原则上，可以应用先前的对称三对角矩阵对角化方法。但是，显式地计算b>b是不明智的，最后一步使用更精细的方法；可以使用问题20.1的矩阵，见问题20.3。同样，见Golub和van Loan[80]，第8章，Demmel[49]和Trefethen和Bau[171]，第31课。

The polar form has applications in continuum mechanics. Indeed, in any deformation it is important to separate stretching from rotation. This is exactly what QS achieves. The orthogonal part Q corresponds to rotation (perhaps with an additional reflection), and the symmetric matrix S to stretching (or compression). The real eigenvalues σ1,...,σr of S are the stretch factors (or compression factors) (see Marsden and Hughes [117]). The fact that S can be diagonalized by an orthogonal matrix corresponds to a natural choice of axes, the principal axes.  
极性形式在连续介质力学中有应用。实际上，在任何变形中，将拉伸与旋转分开都是很重要的。这正是QS所实现的。正交部分q对应于旋转（可能带有附加反射），对称矩阵s对应于拉伸（或压缩）。S的实际特征值σ1，…，σr是拉伸系数（或压缩系数）（见Marsden和Hughes[117]）。事实上，S可以由一个正交矩阵对角化对应于轴的自然选择，主轴。

The SVD has applications to data compression, for instance in image processing. The idea is to retain only singular values whose magnitudes are significant enough. The SVD can also be used to determine the rank of a matrix when other methods such as Gaussian elimination produce very small pivots. One of the main applications of the SVD is the computation of the pseudo-inverse. Pseudo-inverses are the key to the solution of various optimization problems, in particular the method of least squares. This topic is discussed in the next chapter (Chapter 21). Applications of the material of this chapter can be found in Strang [165, 164]; Ciarlet [41]; Golub and Van Loan [80], which contains many other references; Demmel [49]; and Trefethen and Bau [171].  
SVD应用于数据压缩，例如图像处理。其思想是只保留数量足够大的奇异值。当其他方法如高斯消去产生非常小的支点时，SVD也可以用来确定矩阵的秩。支持向量机的主要应用之一是伪逆的计算。伪逆是求解各种优化问题的关键，尤其是最小二乘法。本主题将在下一章（第21章）中讨论。本章材料的应用可在Strang[165164]、Ciarlet[41]、Golub和van Loan[80]中找到，其中包含许多其他参考文献；Demmel[49]、Trefethen和Bau[171]。

## 20.5 Ky Fan Norms and Schatten Norms 20.5 KY Fan规范和Schatten规范

The singular values of a matrix can be used to define various norms on matrices which have found recent applications in quantum information theory and in spectral graph theory. Following Horn and Johnson [93] (Section 3.4) we can make the following definitions:  
矩阵的奇异值可以用来定义矩阵上的各种范数，这些范数在量子信息理论和谱图理论中有着最新的应用。根据Horn和Johnson[93]（第3.4节），我们可以做出以下定义：

Definition 20.5. For any matrix A ∈ Mm,n(C), let q = min{m,n}, and if σ1 ≥ ··· ≥ σq are the singular values of A, for any k with 1 ≤ k ≤ q, let  
定义20.5.对于任意矩阵a∈m m，n（c），设q=min m，n，如果σ1≥······································

Nk(A) = σ1 + ··· + σk,  
nk（a）=σ1+·····+σk，

called the Ky Fan k-norm of A.  
称为A的KY范k-范数。

More generally, for any p ≥ 1 and any k with 1 ≤ k ≤ q, let  
更一般地说，对于任何p≥1和任何k（1≤k≤q），让

,  
，

called the Ky Fan p-k-norm of A. When k = q, Nq;p is also called the Schatten p-norm.  
称为a的ky fan p-k-norm。当k=q，nq；p也称为schatten p-norm。

Observe that when k = 1, N1(A) = σ1, and the Ky Fan norm N1 is simply the spectral norm from Chapter 8, which is the subordinate matrix norm associated with the Euclidean norm. When k = q, the Ky Fan norm Nq is given by  
注意，当k=1时，n1（a）=σ1，而ky-fan范数n1只是第8章中的谱范数，它是与欧几里得范数相关联的次矩阵范数。当k=q时，Ky扇范数nq由下式给出：

Nq(A) = σ1 + ··· + σq = tr((A∗A)1/2)  
nq（a）=σ1+·····+σq=tr（（a a）1/2）

### 20.6. SUMMARY 20.6。总结

and is called the trace norm or nuclear norm. When p = 2 and k = q, the Ky Fan Nq;2 norm is given by  
称为痕迹规范或核规范。当p=2和k=q时，ky fan nq；2范数由下式给出

,  
，

which is the Frobenius norm of A.  
这是A的Frobenius规范。

It can be shown that Nk and Nk;p are unitarily invariant norms, and that when m = n, they are matrix norms; see Horn and Johnson [93] (Section 3.4, Corollary 3.4.4 and Problem  
可以看出，nk和nk；p是统一不变范数，当m=n时，它们是矩阵范数；见Horn和Johnson[93]（第3.4节，推论3.4.4和问题

3).  
3）。

## 20.6 Summary 20.6总结

The main concepts and results of this chapter are listed below:  
本章的主要概念和结果如下：

* For any linear map f : E → E on a Euclidean space E, the maps f∗ ◦f and f ◦f∗ are self-adjoint and positive semidefinite.  
  对于欧几里得空间E上的任何线性映射f:e→e，该映射f f和f f是自伴半定的。
* The singular values of a linear map.  
  线性映射的奇异值。
* Positive semidefinite and positive definite self-adjoint maps.  
  正半定和正定自伴映射。
* Relationships between Imf, Kerf, Imf∗, and Kerf∗.  
  imf、kerf、imf和kerf之间的关系。
* The singular value decomposition theorem for square matrices (Theorem 20.3).  
  平方矩阵的奇异值分解定理（定理20.3）。
* The SVD of matrix.  
  矩阵的SVD。
* The polar decomposition of a matrix.  
  矩阵的极分解。
* The Weyl inequalities.  
  韦尔不等式。
* The singular value decomposition theorem for m × n matrices (Theorem 20.5).  
  m×n矩阵的奇异值分解定理（定理20.5）。
* Ky Fan k-norms, Ky Fan p-k-norms, Schatten p-norms.  
  KY-Fan k-规范，KY-Fan p-k-规范，Schatten p-规范。

## 20.7 Problems 20.7问题

Problem 20.1. (1) Let A be a real n×n matrix and consider the (2n)×(2n) real symmetric matrix  
问题20.1。（1）设a为实n×n矩阵，并考虑（2n）×（2n）实对称矩阵

.  
.

Suppose that A has rank r. If A = V ΣU> is an SVD for A, with Σ = diag(σ1,...,σn) and σ1 ≥ ··· ≥ σr > 0, denoting the columns of U by uk and the columns of V by vk, prove that σk is an eigenvalue of S with corresponding eigenvector , and that −σk is an eigenvalue of S with corresponding eigenvector .  
假设a的秩为r，如果a=v∑u>是a的SVD，用∑=diag（σ1，…，σn）和σ1≥·······················································Nding特征向量。

Hint. We have Auk = σkvk for k = 1,...,n. Show that A>vk = σkuk for k = 1,...,r, and that A>vk = 0 for k = r + 1,...,n. Recall that Ker(A>) = Ker(AA>).  
暗示。对于k=1，…，n，我们有auk=σk vk。表明，对于k=1，…，r，a>vk=σkuk，而对于k=r+1，…，n，a>vk=0。回想一下，ker（a>）=ker（aa>）。

1. Prove that the 2n eigenvectors of S in (1) are pairwise orthogonal. Check that if A has rank r, then S has rank 2r.  
   证明（1）中s的2n特征向量是成对正交的。检查A是否具有等级R，则S是否具有等级2R。
2. Now assume that A is a real m × n matrix and consider the (m + n) × (m + n) real symmetric matrix  
   假设A是实M×N矩阵，考虑（M+N）×实对称矩阵

.  
.

Suppose that A has rank r. If A = V ΣU> is an SVD for A, prove that σk is an eigenvalue of S with corresponding eigenvector , and that −σk is an eigenvalue of  
假设a有秩r，如果a=v∑u>是a的svd，证明σk是具有相应特征向量的s的特征值，而−σk是

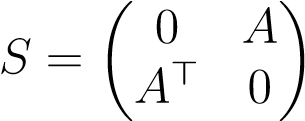
S with corresponding eigenvector .  
与相应的特征向量。

Find the remaining m + n − 2r eigenvectors of S associated with the eigenvalue 0.  
找到与特征值0相关的s的剩余m+n−2r特征向量。

(4) Prove that these m + n eigenvectors of S are pairwise orthogonal.  
（4）证明了S的M+N特征向量是成对正交的。

Problem 20.2. Let A be a real m × n matrix of rank r and let q = min(m,n).  
问题20.2。设a为秩r的实m×n矩阵，设q=min（m，n）。

1. Consider the (m + n) × (m + n) real symmetric matrix  
   考虑（m+n）×（m+n）实对称矩阵



and prove that  
并证明

.  
.

Use the above equation to prove that  
用上述方程证明

det(zIm+n − S) = tn−m det(t2Im − AA>).  
det（zim+n−s）=tn−m det（t2im−aa>）。

1. Prove that the eigenvalues of S are ±σ1,...,±σq, with |m − n| additional zeros.  
   证明了S的特征值是±σ1，…，±σq，加上m−n 0。

Problem 20.3. Let B be a real bidiagonal matrix of the form  
问题20.3。设b为形式的实双对角矩阵

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| a1 网络错误   0 网络错误   网络错误  B =  ... 网络错误    网络错误   0 网络错误  0 网络错误 | b1 网络错误  a2 网络错误  ... ··· 网络错误  0 网络错误 | 0 网络错误  b2 网络错误  ... 网络错误  0 网络错误  ··· 网络错误 | ··· 网络错误  ... 网络错误  ... 网络错误  an−1 网络错误  0 网络错误 | 0  网络错误  0  网络错误  ... . 网络错误  bn−1 网络错误  an 网络错误 |

### 20.7. PROBLEMS 20.7。问题

Let A be the (2n) × (2n) symmetric matrix  
设A为（2n）×（2n）对称矩阵

,  
，

and let P be the permutation matrix given by P = [e1,en+1,e2,en+2,··· ,en,e2n].  
设p为p=[e1，en+1，e2，en+2，···，en，e2n]给出的置换矩阵。

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| diagonal of the form 网络错误 |  |  |  |  |  |  | |
|  网络错误  0 网络错误  a1 网络错误    0 网络错误   网络错误   0 网络错误  T =  ... 网络错误   网络错误    0   0 网络错误   网络错误  0 网络错误 | a1 0 网络错误  b1 网络错误  0 网络错误  ... 网络错误  0 网络错误  0 网络错误  0 网络错误 | 0 网络错误  b1 0 a2 网络错误  ... 网络错误  0 网络错误  0 网络错误  0 网络错误 | 0 0 a2 网络错误  0 网络错误  ... 网络错误  ··· 网络错误  ··· 网络错误  ··· 网络错误 | 0 0 网络错误  0 网络错误  b2 网络错误  ... 网络错误  an−1 网络错误  0 网络错误  0 网络错误 | 0 0 网络错误  0 0 网络错误  ... 网络错误  0 网络错误  bn−1 网络错误  0 网络错误 | ··· 网络错误  ··· ··· 网络错误  ···... 网络错误  bn−1 0 an 网络错误 | 0  网络错误  0  网络错误  0  网络错误  0  网络错误  ... . 网络错误  0  an 网络错误  0 网络错误 |

1. Prove that T = P >AP is a symmetric tridiagonal (2n)×(2n) matrix with zero main  
   证明t=p>a p是一个主零点的对称三对角（2n）×2n矩阵
2. Prove that if xi is a unit eigenvector for an eigenvalue λi of T, then λi = ±σi where σi is a singular value of B, and that  
   证明了如果Xi是T的特征值Li i的单位特征向量，则Li I＝±Sigi I，其中Sigi I是B的奇异值，并且

,  
，

where the ui are unit eigenvectors of B>B and the vi are unit eigenvectors of BB>. Problem 20.4. Find the SVD of the matrix  
其中，ui是b>b的单位特征向量，vi是bb>的单位特征向量。问题20.4。找到矩阵的SVD

.  
.

Problem 20.5. Let u,v ∈ Rn be two nonzero vectors, and let A = uv> be the corresponding rank 1 matrix. Prove that the nonzero singular value of A is kuk2 kvk2.  
问题20.5。设u，v∈rn为两个非零向量，a=uv>为相应的秩1矩阵。证明了A的非零奇异值是kuk2 kvk2。

Problem 20.6. Let A be a n×n real matrix. Prove that if σ1,...,σn are the singular values of A, thenare the singular values of AA>A.  
问题20.6。设a为n×n实矩阵。证明如果σ1，…，σn是a的奇异值，那么aa>a的奇异值。

Problem 20.7. Let A be a real n × n matrix.  
问题20.7。设A为实n×n矩阵。

1. Prove that the largest singular value σ1 of A is given by  
   证明a的最大奇异值σ1由下式给出

,  
，

and that this supremum is achieved at x = u1, the first column in U in an SVD A = V ΣU>.  
这个上确界是在x=u1处得到的，在svd a=v∑u>中u的第一列。

1. Extend the above result to real m × n matrices.  
   将上述结果扩展到实M×N矩阵。

Problem 20.8. Let A be a real m × n matrix. Prove that if B is any submatrix of A (by keeping M ≤ m rows and N ≤ n columns of A), then (σ1)B ≤ (σ1)A (where (σ1)A is the largest singular value of A and similarly for (σ1)B).  
问题20.8。设A为实M×N矩阵。证明如果b是a的任何子矩阵（通过保持m≤m行，n≤n列a），则（σ1）b≤（σ1）a（其中（σ1）a是a的最大奇异值，与（σ1）b相似）。

Problem 20.9. Let A be a real n × n matrix.  
问题20.9。设A为实n×n矩阵。

1. Assume A is invertible. Prove that if A = Q1S1 = Q2S2 are two polar decompositions of A, then Q1 = Q2 and S1 = S2.  
   假设a是可逆的。证明如果a=q1 s1=q2 s2是a的两个极分解，那么q1=q2，s1=s2。

Hint. , with S1 and S2 symmetric positive definite. Then use Problem 16.7.  
暗示。，具有s1和s2对称正定。然后使用问题16.7。

1. Now assume that A is singular. Prove that if A = Q1S1 = Q2S2 are two polar decompositions of A, then S1 = S2, but Q1 may not be equal to Q2.  
   现在假设a是单数。证明如果a=q1 s1=q2 s2是a的两个极分解，那么s1=s2，但q1可能不等于q2。

Problem 20.10. (1) Let A be any invertible (real) n × n matrix. Prove that for every SVD, A = V DU> of A, the product V U> is the same (i.e., if , then  
问题20.10。（1）设A为任意可逆（实）n×n矩阵。证明对于每个SVD，a=v du>a，产品v u>是相同的（即，如果，那么

). What does V U> have to do with the polar form of A?  
）v u>与a的极性形式有什么关系？

(2) Given any invertible (real) n × n matrix, A, prove that there is a unique orthogonal matrix, Q ∈ O(n), such that kA − QkF is minimal (under the Frobenius norm). In fact, prove that Q = V U>, where A = V DU> is an SVD of A. Moreover, if det(A) > 0, show that Q ∈ SO(n).  
（2）给定任意可逆（实）n×n矩阵，a，证明存在唯一的正交矩阵q∈o（n），使得ka−qkf最小（在frobenius范数下）。事实上，证明了q=v u>，其中a=v du>是a的一个svd，并且，如果det（a）>0，则表明q∈so（n）。

What can you say if A is singular (i.e., non-invertible)?  
如果a是单数（即不可逆），你能说什么？

Problem 20.11. (1) Prove that for any n × n matrix A and any orthogonal matrix Q, we have max{tr(QA) | Q ∈ O(n)} = σ1 + ··· + σn,  
问题20.11。（1）证明对于任意n×n矩阵a和任意正交矩阵q，我们有max tr（qa）q∈o（n）=σ1+······+σn，

where σ1 ≥ ··· ≥ σn are the singular values of A. Furthermore, this maximum is achieved by Q = UV >, where A = V ΣU> is any SVD for A.  
其中，σ1≥··············································

(2) By applying the above result with A = Z>X and Q = R>, deduce the following result : For any two fixed n × k matrices X and Z, the minimum of the set  
（2）通过将上述结果与a=z>x和q=r>一起应用，推导出以下结果：对于任意两个固定的n×k矩阵x和z，集合的最小值

{kX − ZRkF | R ∈ O(k)}  
kx−zrkf r∈o（k）

is achieved by R = V U> for any SVD decomposition V ΣU> = Z>X of Z>X.  
对于任何SVD分解V∑U>=Z>X，通过R=V U>实现。

Remark: The problem of finding an orthogonal matrix R such that ZR comes as close as possible to X is called the orthogonal Procrustes problem; see Strang [166] (Section IV.9) for the history of this problem.  
注：找到一个正交矩阵r，使zr尽可能接近x的问题称为正交procrustes问题；关于这个问题的历史，见strang[166]（第4.9节）。

Chapter 21  
第二十一章

# Applications of SVD and Pseudo-Inverses 支持向量机和伪逆的应用

De tous les principes qu’on peut proposer pour cet objet, je pense qu’il n’en est pas de plus g´en´eral, de plus exact, ni d’une application plus facile, que celui dont nous avons fait usage dans les recherches pr´ec´edentes, et qui consiste `a rendre minimum la somme des carr´es des erreurs. Par ce moyen il s’´etablit entre les erreurs une sorte d’´equilibre qui, empˆechant les extrˆemes de pr´evaloir, est tr`es propre `as faire connaitre l’´etat du syst`eme le plus proche de la v´erit´e.  
我们的原则是，我们的建议者必须保证所有的工作都能顺利完成，而且必须准确无误，申请程序简单易行，我们不知道如何使用这些资源，并考虑到至少要有一个完整的工作流程。瑞尔。《欧洲货币体系》中的“电子商务中心”未经分类的“公平报价”，《欧洲货币体系外部评估》，《欧洲货币体系》中的“公平竞争”部分。

—Legendre, 1805, Nouvelles M´ethodes pour la d´etermination des Orbites des Com`etes  
-勒让德，1805年，《轨道测量的新方法》

## 21.1 Least Squares Problems and the Pseudo-Inverse 21.1最小二乘问题和伪逆问题

This chapter presents several applications of SVD. The first one is the pseudo-inverse, which plays a crucial role in solving linear systems by the method of least squares. The second application is data compression. The third application is principal component analysis (PCA), whose purpose is to identify patterns in data and understand the variance–covariance structure of the data. The fourth application is the best affine approximation of a set of data, a problem closely related to PCA.  
本章介绍了SVD的几种应用。第一种是伪逆，它在最小二乘法求解线性系统中起着至关重要的作用。第二个应用程序是数据压缩。第三个应用是主成分分析（PCA），其目的是识别数据中的模式并了解数据的方差-协方差结构。第四个应用是一组数据的最佳仿射近似，这是一个与主成分分析密切相关的问题。

The method of least squares is a way of “solving” an overdetermined system of linear equations  
最小二乘法是求解超定线性方程组的一种方法。

Ax = b,  
ax=b，

i.e., a system in which A is a rectangular m×n matrix with more equations than unknowns (when m > n). Historically, the method of least squares was used by Gauss and Legendre to solve problems in astronomy and geodesy. The method was first published by Legendre in 1805 in a paper on methods for determining the orbits of comets. However, Gauss had already used the method of least squares as early as 1801 to determine the orbit of the asteroid  
也就是说，其中a是一个矩形m×n矩阵，方程多于未知方程（当m>n时）。历史上，高斯和勒让德使用最小二乘法来解决天文学和大地测量学中的问题。勒让德于1805年在一篇关于彗星轨道确定方法的论文中首次发表了这一方法。然而，早在1801年，高斯就已经使用最小二乘法来确定小行星的轨道。

631  
六百三十一

Ceres, and he published a paper about it in 1810 after the discovery of the asteroid Pallas. Incidentally, it is in that same paper that Gaussian elimination using pivots is introduced.  
在发现小行星帕拉斯之后，他于1810年发表了一篇关于它的论文。顺便说一句，在同一篇文章中，介绍了利用支点进行高斯消元。

The reason why more equations than unknowns arise in such problems is that repeated measurements are taken to minimize errors. This produces an overdetermined and often inconsistent system of linear equations. For example, Gauss solved a system of eleven equations in six unknowns to determine the orbit of the asteroid Pallas.  
在这些问题中，产生的方程多于未知数的原因是为了使误差最小化而重复测量。这就产生了一个超定的，经常不一致的线性方程组。例如，高斯在六个未知数中解出了一个由十一个方程组成的系统来确定小行星帕拉斯的轨道。

Example 21.1. As a concrete illustration, suppose that we observe the motion of a small object, assimilated to a point, in the plane. From our observations, we suspect that this point moves along a straight line, say of equation y = dx + c. Suppose that we observed the moving point at three different locations (x1,y1), (x2,y2), and (x3,y3). Then we should have  
例21.1。作为一个具体的例子，假设我们观察一个小物体的运动，被同化到平面上的一个点上。根据我们的观察，我们怀疑这一点沿着直线移动，例如方程y=dx+c。假设我们在三个不同的位置（x1、y1）、（x2、y2）和（x3、y3）观察到移动点。那么我们应该

c + dx1 = y1, c + dx2 = y2, c + dx3 = y3.  
C+dx1=y1，C+dx2=y2，C+dx3=y3。

If there were no errors in our measurements, these equations would be compatible, and c and d would be determined by only two of the equations. However, in the presence of errors, the system may be inconsistent. Yet we would like to find c and d!  
如果在我们的测量中没有误差，这些方程将是兼容的，C和D将只由两个方程确定。但是，在出现错误时，系统可能不一致。但是我们想找到C和D！

The idea of the method of least squares is to determine (c,d) such that it minimizes the sum of the squares of the errors, namely,  
最小二乘法的思想是确定（c，d），使误差平方和最小化，即，

(c + dx1 − y1)2 + (c + dx2 − y2)2 + (c + dx3 − y3)2.  
（C+DX1−Y1）2+（C+DX2−Y2）2+（C+DX3−Y3）2.

See Figure 21.1.  
见图21.1。

y

=

c

x

+

d

(

)

x , y

1

1

)

x , y

(

2

2

x , y

)

(

3

3

1

(

x , cx +d

)

1

)

x , cx +d

(

x , cx +d

)

(

2

2

3

3

)

(

x , y

1

1

(

x , y

)

2

2

(

x , y

)

3

3

Figure 21.1: Given three points (x1,y1), (x2,y2), (x3,y3), we want to determine the line y = cx + d which minimizes the lengths of the dashed vertical lines.  
图21.1：给定三个点（x1，y1），（x2，y2），（x3，y3），我们想确定一条线y=cx+d，它将虚线垂直线的长度最小化。

In general, for an overdetermined m × n system Ax = b, what Gauss and Legendre discovered is that there are solutions x minimizing  
一般来说，对于一个超定的m×n系统ax=b，高斯和勒让德发现有解x最小化



(where, the square of the Euclidean norm of the vector u = (u1,...,un)), and that these solutions are given by the square n × n system  
（式中，向量u的欧几里得范数的平方=（u1，…，u n）），这些解由平方n×n系统给出。

A>Ax = A>b,  
a>ax=a>b，

called the normal equations. Furthermore, when the columns of A are linearly independent, it turns out that A>A is invertible, and so x is unique and given by  
称为法方程。此外，当a的列是线性无关的时，结果表明a>a是可逆的，因此x是唯一的，由

x = (A>A)−1A>b.  
x=（a>a）−1a>b.

Note that A>A is a symmetric matrix, one of the nice features of the normal equations of a least squares problem. For instance, since the above problem in matrix form is represented as  
注意a>a是一个对称矩阵，是最小二乘法方程的一个很好的特征。例如，由于上述矩阵形式的问题表示为

,  
，

the normal equations are  
正态方程是

.  
.

In fact, given any real m × n matrix A, there is always a unique x+ of minimum norm that minimizes, even when the columns of A are linearly dependent. How do we prove this, and how do we find x+?  
事实上，对于任何实M×N矩阵A，总是有一个唯一的最小范数x+，即使A的列是线性相关的，也会最小化。我们如何证明这一点，以及如何找到x+？

Theorem 21.1. Every linear system Ax = b, where A is an m × n matrix, has a unique least squares solution x+ of smallest norm.  
定理21.1。每个线性系统a x=b，其中a是m×n矩阵，具有最小范数的唯一最小二乘解x+。

Proof. Geometry offers a nice proof of the existence and uniqueness of x+. Indeed, we can interpret b as a point in the Euclidean (affine) space Rm, and the image subspace of A (also called the column space of A) as a subspace U of Rm (passing through the origin). Then it is clear that  
证据。几何为X+的存在和唯一性提供了很好的证明。实际上，我们可以把b解释为欧几里得空间rm中的一个点，把a的图像子空间（也称为a的列空间）解释为rm的子空间u（通过原点）。很明显

,  
，

with U = ImA, and we claim that x minimizes kAx−, where p the orthogonal projection of b onto the subspace U.  
当u=ima时，我们认为x使kax-最小化，其中p是b在子空间u上的正交投影。

Recall from Section 12.1 that the orthogonal projection pU : U ⊕ U⊥ → U is the linear map given by  
从第12.1节回忆起，正交投影pu:u u→u是由

pU(u + v) = u,  
pu（u+v）=u，

with u ∈ U and v ∈→−U⊥. If we let− ∈ p = pU(b) ∈ U, then for any point y ∈ U, the vectors py−→ = y − p ∈ U and bp = p b U⊥ are orthogonal, which implies that  
带u∈u和v∈→−u。如果我们让−p=p u（b）∈u，那么对于任意点y∈u，向量py−→=y−p∈u和bp=p b u是正交的，这意味着

,  
，

where →−by = y −b. Thus, p is indeed the unique point in U that minimizes the distance from b to any point in U. See Figure 21.2.  
其中→−b y=y−b。因此，p确实是u中的唯一点，它将b到u中任何点的距离减至最小。见图21.2。

Im A = U

b

p

Im A = U

b

p

y

Figure 21.2: Given a 3 × 2 matrix A, U = ImA is the peach plane in R3 and p is the orthogonal projection of b onto U. Furthermore, given y ∈ U, the points b, y, and p are the vertices of a right triangle.  
图21.2：给定3×2矩阵a，u=ima是r3中的桃平面，p是b在u上的正交投影。此外，给定y∈u，点b、y和p是直角三角形的顶点。

Thus the problem has been reduced to proving that there is a unique x+ of minimum norm such that Ax+ = p, with p = pU(b) ∈ U, the orthogonal projection of b onto U. We use the fact that  
因此，该问题被简化为证明存在一个唯一的最小范数x+，使得ax+=p，其中p=p u（b）∈u，b对u的正交投影。

Rn = KerA ⊕ (KerA)⊥.  
RN=KERA（KERA）。

Consequently, every x ∈ Rn can be written uniquely as x = u + v, where u ∈ KerA and v ∈ (KerA)⊥, and since u and v are orthogonal,  
因此，每个x∈rn可以唯一地写成x=u+v，其中u∈kera和v∈（kera），并且由于u和v是正交的，

.  
.

Furthermore, since u ∈ KerA, we have Au = 0, and thus Ax = p iff Av = p, which shows that the solutions of Ax = p for which x has minimum norm must belong to (KerA)⊥. However, the restriction of A to (KerA)⊥ is injective. This is because if Av1 = Av2, where v1,v2 ∈ (KerA)⊥, then A(v2 − v2) = 0, which implies v2 − v1 ∈ KerA, and since v1,v2 ∈ (KerA)⊥, we also have v2 − v1 ∈ (KerA)⊥, and consequently, v2 − v1 = 0. This shows that there is a unique x+ of minimum norm such that Ax+ = p, and that x+ must belong to  
此外，由于u∈kera，我们得到了au=0，因此ax=p iff av=p，这表明x具有最小范数的ax=p的解必须属于（kera）。但是，A到（KERA）的限制是注射性的。这是因为如果a v1=a v2，其中v1，v2∈（kera），那么a（v2−v2）=0，这意味着v2−v1∈kera，并且由于v1，v2∈（kera），我们也有v2−v1∈（kera），因此，v2−v1=0。这表明有一个最小范数的唯一x+，这样ax+=p，x+必须属于

(KerA)⊥. By our previous reasoning, x+ is the unique vector of minimum norm minimizing  
（喀拉邦）。根据前面的推理，x+是最小范数最小的唯一向量。

.   
.

The proof also shows that x minimizesis orthogonal to U, which can be expressed by saying that b−Ax is orthogonal to every column of A. However, this is equivalent to  
证明还表明，x与u正交，可以用b-ax与a的每一列正交来表示。然而，这相当于

A>(b − Ax) = 0, i.e., A>Ax = A>b.  
a>（b−ax）=0，即a>ax=a>b。

Finally, it turns out that the minimum norm least squares solution x+ can be found in terms of the pseudo-inverse A+ of A, which is itself obtained from any SVD of A.  
最后，证明了最小范数最小二乘解x+可以用a的伪逆a+来表示，a的伪逆a+本身就是从a的任意svd得到的。

Definition 21.1. Given any nonzero m × n matrix A of rank r, if A = V DU> is an SVD of A such that  
定义21.1.给定秩r的任何非零m×n矩阵a，如果a=v du>是a的svd，

,  
，

with  
具有

Λ = diag(λ1,...,λr)  
∧=diag（λ1，…，λr）

an r ×r diagonal matrix consisting of the nonzero singular values of A, then if we let D+ be the n × m matrix  
一个由a的非零奇异值组成的r×r对角矩阵，如果我们让d+是n×m矩阵

,  
，

with  
具有

Λ−1 = diag(1/λ1,...,1/λr),  
∧−1=diag（1/λ1，…，1/λr），

the pseudo-inverse of A is defined by  
a的伪逆定义为

A+ = UD+V >.  
a+=ud+v>。

If A = 0m,n is the zero matrix, we set A+ = 0n,m. Observe that D+ is obtained from D by inverting the nonzero diagonal entries of D, leaving all zeros in place, and then transposing the matrix. For example, given the matrix  
如果a=0 m，n是零矩阵，我们设置a+=0 n，m。观察d+是通过颠倒d的非零对角线项得到的，保留所有零，然后转置矩阵。例如，给定矩阵

,  
，

its pseudo-inverse is  
它的伪逆是

The pseudo-inverse of a matrix is also known as the Moore–Penrose pseudo-inverse.  
矩阵的伪逆矩阵也称为摩尔-彭罗斯伪逆矩阵。

Actually, it seems that A+ depends on the specific choice of U and V in an SVD (U,D,V ) for A, but the next theorem shows that this is not so.  
实际上，a+似乎取决于a的svd（u，d，v）中u和v的具体选择，但是下一个定理表明这不是这样的。

Theorem 21.2. The least squares solution of smallest norm of the linear system Ax = b, where A is an m × n matrix, is given by  
定理21.2。线性系统最小范数的最小二乘解ax=b，其中a是m×n矩阵，由下式得出：

x+ = A+b = UD+V >b.  
x+=a+b=ud+v>b。

Proof. First assume that A is a (rectangular) diagonal matrix D, as above. Then since x minimizesis the projection of b onto the image subspace F of D, it is fairly obvious that x+ = D+b. Otherwise, we can write  
证据。首先假设a是（矩形）对角矩阵d，如上所述。然后，由于x最小化了b在d的图像子空间f上的投影，很明显x+=d+b。否则，我们可以写

A = V DU>,  
a=v du>，

where U and V are orthogonal. However, since V is an isometry,  
其中u和v是正交的。但是，既然v是等距测量，

kAx − bk2 = kV DU>x − bk2 = kDU>x − V >bk2.  
kax−bk2=kv du>x−bk2=kdu>x−v>bk2。

Letting y = U>x, we have kxk2 = kyk2, since U is an isometry, and since U is surjective, kAx − bk2 is minimized iff kDy − V >bk2 is minimized, and we have shown that the least solution is  
假设y=u>x，我们得到kxk2=kkk2，因为u是一个等值线，并且由于u是可预测的，所以当kdy−v>bk2最小化时，kax−bk2最小化，并且我们已经证明了最小解是

y+ = D+V >b.  
Y+=D+V>B。

Since y = U>x, with kxk2 = kyk2, we get  
因为y=u>x，kxk2=kkk2，我们得到

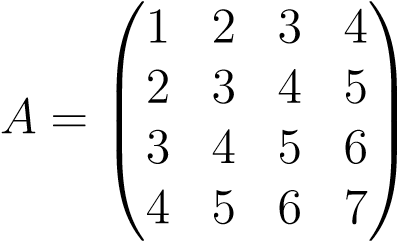
x+ = UD+V >b = A+b.  
x+=ud+v>b=a+b。

Thus, the pseudo-inverse provides the optimal solution to the least squares problem.   
因此，伪逆为最小二乘问题提供了最优解。

By Theorem 21.2 and Theorem 21.1, A+b is uniquely defined by every b, and thus A+ depends only on A.  
根据定理21.2和21.1，a+b由每个b唯一定义，因此a+仅依赖于a。

The Matlab command for computing the pseudo-inverse B of the matrix A is B = pinv(A).  
用于计算矩阵A的伪逆B的matlab命令是b=pinv（a）。

Example 21.2. If A is the rank 2 matrix  
例21.2。如果a是秩2矩阵



whose eigenvalues are −1.1652,0,0,17.1652, using Matlab we obtain the SVD A = V DU>  
其特征值为−1.1652,0,0,17.1652，使用matlab，我们得到了SVD a=v du>

with  
具有

|  |  |  |  |
| --- | --- | --- | --- |
| −0.3147 网络错误   1. = −−00..42755402 网络错误    网络错误  −0.6530 网络错误  −0.3147 网络错误   1. = −−00..42755402 网络错误    网络错误  −0.6530 网络错误 | 0.7752 网络错误  0.3424 网络错误  −00..09035231 网络错误  − 网络错误  −00..77523424 网络错误  −0.0903 网络错误  0.5231 网络错误 | 0.2630 网络错误  0.0075 网络错误  −0.8039 网络错误  0.5334 网络错误  0.5452 网络错误  −0.7658 网络错误  −0.1042 网络错误  0.3247 网络错误 | −0.4805 网络错误  0.8366  网络错误  −0.2319, 网络错误  −0.1243 网络错误  0.0520  网络错误  0.3371  网络错误  −0.8301, 网络错误  0.4411 网络错误 |

17.1652 0 0 0  
17.1652 0 0\_

D =  00 1.16520 00 00.  
D=00 1.16520 00 00。

0 0 0 0  
0 0 0 0 0

Then  
然后

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| and 网络错误 |  | 0 网络错误  0.8583 0 网络错误  0 网络错误 | 0 0 网络错误  0 0, 网络错误  0 0 网络错误  0 0 网络错误 |  |
|  |  | −00..22000900 网络错误  −0.0400 网络错误  0.1700 网络错误 | 0.0700 0.0400 网络错误  0.0100 网络错误  −0.0200 网络错误 | 0.3600  网络错误  0.1700 , −0.0200 网络错误  −0.2100 网络错误 |

which is also the result obtained by calling pinv(A).  
这也是通过调用pinv（a）获得的结果。

If A is an m × n matrix of rank n (and so m ≥ n), it is immediately shown that the  
如果a是n阶的m×n矩阵（因此m≥n），则立即显示

QR-decomposition in terms of Householder transformations applies as follows:  
就户主转换而言，QR分解应用如下：

There are n m × m matrices H1,...,Hn, Householder matrices or the identity, and an upper triangular m × n matrix R of rank n such that  
有n个m×m矩阵h1，…，hn，户主矩阵或恒等式，以及一个上三角m×n矩阵r的秩n，这样

A = H1 ···HnR.  
a=h1···hnr。

Then because each Hi is an isometry,  
因为每个Hi都是一个等距线，

kAx − bk2 = kRx − Hn ···H1bk2,  
kax−bk2=krx−hn···h1bk2，

and the least squares problem Ax = b is equivalent to the system  
最小二乘问题ax=b等于系统

Rx = Hn ···H1b.  
rx=hn···h1b.

Now the system  
现在系统

Rx = Hn H b  
Rx=Hn H b

is of the form  
是这样的

,  
，

where R1 is an invertible n×n matrix (since A has rank n), c ∈ Rn, and d ∈ Rm−n, and the least squares solution of smallest norm is  
其中，R1是可逆n×n矩阵（因为a有秩n），c∈rn，d∈rm−n，最小范数的最小二乘解为

x+ = R1−1c.  
X+=R1−1C。

Since R1 is a triangular matrix, it is very easy to invert R1.  
因为R1是一个三角形矩阵，所以很容易将R1反转。

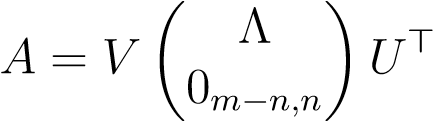
The method of least squares is one of the most effective tools of the mathematical sciences. There are entire books devoted to it. Readers are advised to consult Strang [165], Golub and Van Loan [80], Demmel [49], and Trefethen and Bau [171], where extensions and applications of least squares (such as weighted least squares and recursive least squares) are described. Golub and Van Loan [80] also contains a very extensive bibliography, including a list of books on least squares.  
最小二乘法是数学科学中最有效的工具之一。有很多书都是专门为它写的。建议读者参考Strang[165]、Golub和van Loan[80]、Demmel[49]和Trefethen和Bau[171]，其中描述了最小二乘（如加权最小二乘和递归最小二乘）的扩展和应用。Golub和vanLoan[80]还包含了非常广泛的参考书目，包括关于最小二乘法的书籍列表。

## 21.2 Properties of the Pseudo-Inverse 21.2伪逆函数的性质

We begin this section with a proposition which provides a way to calculate the pseudo-inverse of an m × n matrix A without first determining an SVD factorization.  
我们从一个命题开始，这个命题提供了一种计算m×n矩阵a的伪逆的方法，而不需要首先确定SVD分解。

Proposition 21.3. When A has full rank, the pseudo-inverse A+ can be expressed as A+ = (A>A)−1A> when m ≥ n, and as A+ = A>(AA>)−1 when n ≥ m. In the first case (m ≥ n), observe that A+A = I, so A+ is a left inverse of A; in the second case (n ≥ m), we have AA+ = I, so A+ is a right inverse of A.  
提案21.3.当a具有满秩时，当m≥n时，伪逆a+可表示为a+=（a>a）−1a>，当n≥m时，伪逆a+可表示为a+=a>（aa>）−1。在第一种情况下（m≥n），观察a+a=i，因此a+是a的左逆；在第二种情况下（n≥m），我们有aa+=i，因此a+是a的右逆。

Proof. If m ≥ n and A has full rank n, we have  
证据。如果m≥n且a具有满秩n，则我们有



with Λ an n × n diagonal invertible matrix (with positive entries), so  
带有∧an n×n对角可逆矩阵（带正项），所以

We find that  
我们发现了

,  
，

which yields  
会产生

.  
.

Therefore, if m ≥ n and A has full rank n, then  
因此，如果m≥n且a具有满秩n，则

A+ = (A>A)−1A>. If n ≥ m and A has full rank m, then  
A+=（A>A）−1a>。如果n≥m且a具有满秩m，则



with Λ an m × m diagonal invertible matrix (with positive entries), so  
带∧an m×m对角可逆矩阵（带正项），所以

We find that  
我们发现了

,  
，

which yields  
会产生

.  
.

Therefore, if n ≥ m and A has full rank m, then A+ = A>(AA>)−1.   
因此，如果n≥m且a具有满秩m，则a+=a>（aa>）-1。

### 21.2. PROPERTIES OF THE PSEUDO-INVERSE 21.2。伪逆的性质

For example, if, then A has rank 2 and since m ≥ n, A+ = (A>A)−1A>  
例如，如果，那么a的等级为2，并且由于m≥n，a+=（a>a）−1a>

where  
哪里

.  
.

If, since A has rank 2 and n ≥ m, then A+ = A>(AA>)−1 where  
如果，由于a的秩2和n≥m，则a+=a>（aa>）-1，其中

.  
.

Let A = V ΣU> be an SVD for any m × n matrix A. It is easy to check that both AA+ and A+A are symmetric matrices. In fact,  
假设a=v∑u>是任意m×n矩阵a的SVD，很容易检查aa+和a+a都是对称矩阵。事实上，

and   
和

From the above expressions we immediately deduce that  
从上面的表达式，我们立即推断

|  |  |  |
| --- | --- | --- |
| AA+A 网络错误 | = 网络错误 | A, 网络错误 |
| A+AA+ 网络错误  and that 网络错误 | = 网络错误 | A+, 网络错误 |
| (AA+)2 网络错误 | = 网络错误 | AA+, 网络错误 |
| (A+A)2 网络错误 | = 网络错误 | A+A, 网络错误 |

so both AA+ and A+A are orthogonal projections (since they are both symmetric).  
所以a a+和a+都是正交投影（因为它们都是对称的）。

Proposition 21.4. The matrix AA+ is the orthogonal projection onto the range of A and A+A is the orthogonal projection onto Ker(A)⊥ = Im(A>), the range of A>.  
提案21.4.矩阵a a+是a范围的正交投影，a+a是k（a）=im（a>）的正交投影，a>的范围。

Proof. Obviously, we have range(AA+) ⊆ range(A), and for any y = Ax ∈ range(A), since  
证据。显然，我们有范围（a a+）范围（a），对于任何y=ax∈范围（a），因为

AA+A = A, we have  
a a+a=a，我们有

AA+y = AA+Ax = Ax = y,  
aa+y=aa+ax=ax=y，

so the image of AA+ is indeed the range of A. It is also clear that Ker(A) ⊆ Ker(A+A), and since AA+A = A, we also have Ker(A+A) ⊆ Ker(A), and so  
所以a a+的图像实际上是a的范围，也很明显，ker（a）ker（a+a），由于aa+a=a，我们还有ker（a+a）ker（a），所以

Ker(A+A) = Ker(A).  
ker（A+A）=ker（A）。

Since A+A is symmetric, range(A+A) = range((A+A)>) = Ker(A+A)⊥ = Ker(A)⊥, as claimed.   
由于a+a是对称的，范围（a+a）=range（（a+a）>）=ker（a+a）=ker（a），如权利要求所述。

Proposition 21.5. The set range(A) = range(AA+) consists of all vectors y ∈ Rm such that  
提案21.5。集合范围（a）=range（aa+）由所有向量y∈rm组成，这样

,  
，

with z ∈ Rr.  
带z∈rr。

Proof. Indeed, if y = Ax, then  
证据。实际上，如果y=ax，那么

,  
，

where Σr is the r × r diagonal matrix diag(σ1,...,σr). Conversely, if ), then ), and  
其中∑r是r×r对角矩阵diag（σ1，…，σr）。相反，如果），则），以及



which shows that y belongs to the range of A.   
这表明Y属于A的范围。

Similarly, we have the following result.  
同样，我们得到了以下结果。

Proposition 21.6. The set range(A+A) = Ker(A)⊥ consists of all vectors y ∈ Rn such that  
提案21.6.集合范围（a+a）=ker（a）由所有向量y∈rn组成，这样

,  
，

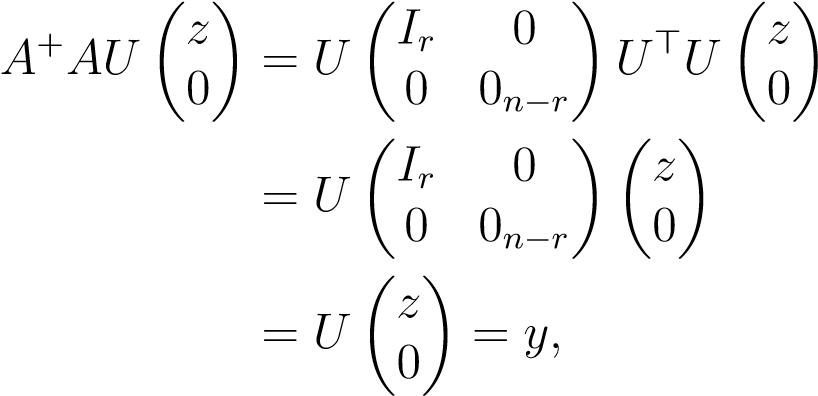
with z ∈ Rr.  
带z∈rr。

### 21.2. PROPERTIES OF THE PSEUDO-INVERSE 21.2。伪逆的性质

Proof. If y = A+Au, then  
证据。如果y=a+au，那么

,  
，

for some z ∈ Rr. Conversely, if), then), and so  
对于某些z∈rr。相反，如果），则），依此类推。



which shows that y ∈ range(A+A).   
表示y∈范围（a+a）。

Analogous results hold for complex matrices, but in this case, V and U are unitary matrices and AA+ and A+A are Hermitian orthogonal projections.  
类似的结果适用于复杂矩阵，但在这种情况下，v和u是一元矩阵，a a+和a+a是厄米特正交投影。

If A is a normal matrix, which means that AA> = A>A, then there is an intimate relationship between SVD’s of A and block diagonalizations of A. As a consequence, the pseudo-inverse of a normal matrix A can be obtained directly from a block diagonalization of A.  
如果a是正态矩阵，即aa>=a>a，则a的svd与a的块对角化之间存在密切关系，因此，正态矩阵a的伪逆可以直接从a的块对角化得到。

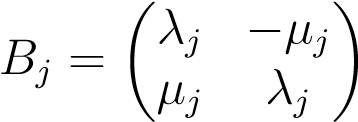
If A is a (real) normal matrix, then we know from Theorem 16.18 that A can be block diagonalized with respect to an orthogonal matrix U as  
如果a是（实数）正规矩阵，那么我们从定理16.18知道a可以对一个正交矩阵u进行分块对角化，作为

A = UΛU>,  
A=U∧U>，

where Λ is the (real) block diagonal matrix  
其中∧是（实数）块对角矩阵

Λ = diag(B1,...,Bn),  
∧=diag（b1，…，bn）

consisting either of 2 × 2 blocks of the form  
由2×2块模板组成



with µj = 06 , or of one-dimensional blocks Bk = (λk). Then we have the following proposition:  
μj=06，或一维块bk=（λk）。然后我们有以下建议：

Proposition 21.7. For any (real) normal matrix A and any block diagonalization A = UΛU> of A as above, the pseudo-inverse of A is given by  
提案21.7.对于任何（实）法向矩阵A和任何块对角化a=u∧u>如上所述，a的伪逆由下式给出：

A+ = UΛ+U>,  
A+=U∧+U>，

where Λ+ is the pseudo-inverse of Λ. Furthermore, if  
其中∧+是∧的伪逆。此外，如果

,  
，

where Λr has rank r, then  
式中，∧r的秩为r，则

.  
.

Proof. Assume that B1,...,Bp are 2×2 blocks and that λ2p+1,...,λn are the scalar entries. We know that the numbers λj ± iµj, and the λ2p+k are the eigenvalues of A. Let ρ2j−1 =  
证据。假设b1，…，bp是2×2块，而λ2p+1，…，λn是标量项。我们知道，数字λj±iμj和λ2p+k是a的特征值，设为ρ2j−1。=

ρ2j = qλ2j + µj2 = pdet(Bi) for j = 1,...,p, ρj = |λj| for j = 2p + 1,...,r. Multiplying U by a suitable permutation matrix, we may assume that the blocks of Λ are ordered so that ρ1 ≥ ρ2 ≥ ··· ≥ ρr > 0. Then it is easy to see that  
对于j=1，…，p，ρj=λj对于j=2p+1，…，r，？2j=qλ2j+μj2=pdet（bi）。将u乘以适当的置换矩阵，我们可以假定∧的块是有序的，因此，ρ1≥ρ2≥·····························那么很容易看出

AA> = A>A = UΛU>UΛ>U> = UΛΛ>U>,  
a a>=a>a=u∧u>u∧>u>=u∧∧>u>，

with  
具有

ΛΛ> = diag(,  
∧∧>=诊断（，

so ρ1 ≥ ρ2 ≥ ··· ≥ ρr > 0 are the singular values σ1 ≥ σ2 ≥ ··· ≥ σr > 0 of A. Define the diagonal matrix  
因此，ρ1≥ρ2≥············································

Σ = diag(σ1,...,σr,0,...,0),  
∑=diag（σ1，…，σr，0，…，0）、

where r = rank(A), σ1 ≥ ··· ≥ σr > 0 and the block diagonal matrix Θ defined such that  
式中，r=秩（a），σ1≥·································

the block Bi in Λ is replaced by the block σ−1Bi where σ = pdet(Bi), the nonzero scalar λj is replaced λj/|λj|, and a diagonal zero is replaced by 1. Observe that Θ is an orthogonal matrix and  
将∧中的块bi替换为块σ−1bi，其中σ=pdet（bi），将非零标量λj替换为λj/λj，将对角零替换为1。观察到θ是一个正交矩阵

Λ = ΘΣ.  
∧=完成∑。

But then we can write  
但是我们可以写

A = UΛU> = UΘΣU>,  
A=U∧U>=U完成∑U>，

and we if let V = UΘ, since U is orthogonal and Θ is also orthogonal, V is also orthogonal and A = V ΣU> is an SVD for A. Now we get  
如果我们让v=u，因为u是正交的，而θ也是正交的，v也是正交的，a=v∑u>是a的svd，现在我们得到

A+ = UΣ+V > = UΣ+Θ>U>.  
A+=U∑+V>=U∑+成人>U>。

However, since Θ is an orthogonal matrix, Θ> = Θ−1, and a simple calculation shows that  
然而，由于θ是一个正交矩阵，所以θ>=θ-1，简单的计算表明

Σ+Θ> = Σ+Θ−1 = Λ+,  
∑+完成>=∑+完成−1=∧+，

which yields the formula  
得出公式

A+ = UΛ+U>.  
A+=U∧+U>。

Also observe that Λr is invertible and  
也注意到∧r是可逆的，并且

.  
.

Therefore, the pseudo-inverse of a normal matrix can be computed directly from any block diagonalization of A, as claimed.   
因此，正态矩阵的伪逆矩阵可以直接从A的任何块对角化中计算出来，如所述。

### 21.3. DATA COMPRESSION AND SVD 21.3。数据压缩和SVD

Example 21.3. Consider the following real diagonal form of the normal matrix  
例21.3。考虑下正规矩阵的实对角形式

,  
，

with  
具有

.  
.

We obtain  
我们得到

,  
，

and the pseudo-inverse of A is  
A的伪逆是

,  
，

which agrees with pinv(A).  
与PINv（a）一致。

The following properties, due to Penrose, characterize the pseudo-inverse of a matrix. We have already proved that the pseudo-inverse satisfies these equations. For a proof of the converse, see Kincaid and Cheney [100].  
由于Penrose的原因，以下特性描述了矩阵的伪逆矩阵。我们已经证明了伪逆满足这些方程。关于相反的证明，见Kincaid和Cheney[100]。

Proposition 21.8. Given any m × n matrix A (real or complex), the pseudo-inverse A+ of A is the unique n × m matrix satisfying the following properties:  
提案21.8。对于任意m×n矩阵a（实矩阵或复矩阵），a的伪逆a+是满足以下特性的唯一n×m矩阵：

AA+A = A,  
a a+a=a，

A+AA+ = A+,  
A+AA+=A+，

(AA+)> = AA+, (A+A)> = A+A.  
（a a+）>=aa+，（a+a）>=a+a。

## 21.3 Data Compression and SVD 21.3数据压缩和SVD

Among the many applications of SVD, a very useful one is data compression, notably for images. In order to make precise the notion of closeness of matrices, we use the notion of matrix norm. This concept is defined in Chapter 8, and the reader may want to review it before reading any further.  
在SVD的众多应用中，一个非常有用的应用是数据压缩，尤其是图像压缩。为了使矩阵的紧密性概念更加精确，我们使用了矩阵范数的概念。这一概念在第8章中有定义，读者可能想在进一步阅读之前回顾一下。

Given an m × n matrix of rank r, we would like to find a best approximation of A by a matrix B of rank k ≤ r (actually, k < r) such that kA − Bk2 (or kA − BkF ) is minimized. The following proposition is known as the Eckart–Young theorem.  
给定秩r的m×n矩阵，我们希望通过秩k≤r的矩阵b（实际上，k<r）找到a的最佳近似值，从而使ka−bk2（或ka−bkf）最小化。下面的命题被称为Eckart-Young定理。

Proposition 21.9. Let A be an m × n matrix of rank r and let V DU> = A be an SVD for A. Write ui for the columns of U, vi for the columns of V , and σ1 ≥ σ2 ≥ ··· ≥ σp for the singular values of A (p = min(m,n)). Then a matrix of rank k < r closest to A (in the k k2 norm) is given by  
提案21.9.设a为秩r的m×n矩阵，v du>=a为a的svd，写出u列的ui，v列的vi，σ1≥σ2≥·······································然后，最接近a（k k2范数）的秩k<r矩阵由下式得出：

diag(σ1,...,σk,0,...,0)U>  
diag（σ1，…，σk，0，…，0）u>

and kA − Akk2 = σk+1.  
kA−akk2=σk+1。

Proof. By construction, Ak has rank k, and we have  
证据。根据结构，AK的等级是K，我们有

p diag(0.  
P诊断（0.

= +1  
= 1

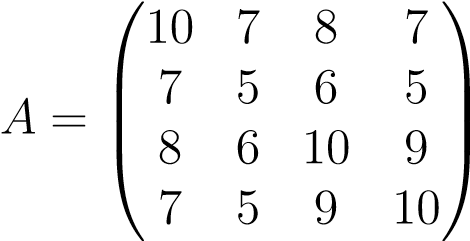
It remains to show that kA − Bk2 ≥ σk+1 for all rank k matrices B. Let B be any rank k matrix, so its kernel has dimension n−k. The subspace Uk+1 spanned by (u1,...,uk+1) has dimension k + 1, and because the sum of the dimensions of the kernel of B and of Uk+1 is (n − k) + k + 1 = n + 1, these two subspaces must intersect in a subspace of dimension at least 1. Pick any unit vector h in Ker(B) ∩ Uk+1. Then since Bh = 0, and since U and V are isometries, we have  
仍然需要证明所有秩k矩阵b的ka−bk2≥σk+1。假设b是任何秩k矩阵，那么它的核具有维数n−k。由（u1，…，uk+1）所跨越的子空间uk+1具有维数k+1，并且因为b和uk+1的核的维数之和是（n−k）+k+1=n+1，t这两个子空间必须在维度至少为1的子空间中相交。选取Ker（b）UK+1中的任何单位向量h。既然bh=0，既然u和v是等距的，我们有

,  
，

which proves our claim.   
这证明了我们的主张。

Note that Ak can be stored using (m + n)k entries, as opposed to mn entries. When , this is a substantial gain.  
请注意，AK可以使用（m+n）k项存储，而不是使用mn项。当，这是一个巨大的收益。

Example 21.4. Consider the badly conditioned symmetric matrix  
例21.4。考虑坏条件对称矩阵



from Section 8.5. Since A is SPD, we have the SVD  
来自第8.5节。既然A是SPD，我们有SVD

A = UDU>,  
A=Udu>，

with  
具有

.  
.

If we set σ3 = σ4 = 0, we obtain the best rank 2 approximation  
如果我们将σ3=σ4=0，我们得到最佳秩2近似值。

.  
.

A nice example of the use of Proposition 21.9 in image compression is given in Demmel [49], Chapter 3, Section 3.2.3, pages 113–115; see the Matlab demo.  
demmel[49]第3章第3.2.3节第113-115页给出了在图像压缩中使用21.9号提案的一个很好的例子；见matlab演示。

Proposition 21.9 also holds for the Frobenius norm; see Problem 21.4.  
提案21.9也适用于弗罗贝尼乌斯规范；见问题21.4。

An interesting topic that we have not addressed is the actual computation of an SVD. This is a very interesting but tricky subject. Most methods reduce the computation of an SVD to the diagonalization of a well-chosen symmetric matrix which is not A>A; see Problem 20.1 and Problem 20.3. Interested readers should read Section 5.4 of Demmel’s excellent book [49], which contains an overview of most known methods and an extensive list of references.  
一个有趣的话题我们还没有讨论，那就是SVD的实际计算。这是一个很有趣但很棘手的问题。大多数方法都将SVD的计算简化为选定的对称矩阵的对角化，该对称矩阵不是a>a；见问题20.1和问题20.3。感兴趣的读者应该阅读德梅尔优秀著作[49]的第5.4节，其中包括最著名的方法概述和大量参考文献。

## 21.4 Principal Components Analysis (PCA) 21.4主要成分分析（PCA）

Suppose we have a set of data consisting of n points X1,...,Xn, with each Xi ∈ Rd viewed as a row vector. Think of the Xi’s as persons, and if Xi = (xi1,...,xid), each xij is the value of some feature (or attribute) of that person.  
假设我们有一组由N点X1，…，Xn组成的数据，每个Xi RD被看作行向量。把XI看成是人，如果X=（XI1，…，XID），每个XIJ都是那个人的某些特征（或属性）的值。

Example 21.5. For example, the Xi’s could be mathematicians, d = 2, and the first component, xi1, of Xi could be the year that Xi was born, and the second component, xi2, the length of the beard of Xi in centimeters. Here is a small data set:  
例21.5。例如，XI可以是数学家，D＝2，XI的第一个组成部分XI1可以是XI出生的一年，第二个组成部分XI2是XI的胡须长度厘米。下面是一个小数据集：

|  |  |  |
| --- | --- | --- |
| Name 网络错误 | year 网络错误 | length 网络错误 |
| Carl Friedrich Gauss 网络错误 | 1777 网络错误 | 0 网络错误 |
| Camille Jordan 网络错误 | 1838 网络错误 | 12 网络错误 |
| Adrien-Marie Legendre 网络错误 | 1752 网络错误 | 0 网络错误 |
| Bernhard Riemann 网络错误 | 1826 网络错误 | 15 网络错误 |
| David Hilbert 网络错误 | 1862 网络错误 | 2 网络错误 |
| Henri Poincar´e 网络错误 | 1854 网络错误 | 5 网络错误 |
| Emmy Noether 网络错误 | 1882 网络错误 | 0 网络错误 |
| Karl Weierstrass 网络错误 | 1815 网络错误 | 0 网络错误 |
| Eugenio Beltrami 网络错误 | 1835 网络错误 | 2 网络错误 |
| Hermann Schwarz 网络错误 | 1843 网络错误 | 20 网络错误 |

We usually form the n × d matrix X whose ith row is Xi, with 1 ≤ i ≤ n. Then the jth column is denoted by Cj (1 ≤ j ≤ d). It is sometimes called a feature vector, but this terminology is far from being universally accepted. In fact, many people in computer vision call the data points Xi feature vectors!  
我们通常形成n×d矩阵x，其行是Xi，1的i i小于n，然后用Cj（1×j j）表示JTH列。它有时被称为特征向量，但这个术语远未被普遍接受。其实很多人在计算机视觉上调用了数据点XI的特征向量！

The purpose of principal components analysis, for short PCA, is to identify patterns in data and understand the variance–covariance structure of the data. This is useful for the following tasks:  
主成分分析（简称PCA）的目的是识别数据模式，了解数据的方差-协方差结构。这对以下任务很有用：

1. Data reduction: Often much of the variabi lity of the data can be accounted for by a smaller number of principal components.  
   数据简化：通常，数据的许多变量都可以由较少的主成分来解释。
2. Interpretation: PCA can show relationships that were not previously suspected.  
   解释：PCA可以显示以前没有被怀疑的关系。

Given a vector (a sample of measurements) x = (x1,...,xn) ∈ Rn, recall that the mean (or average) x of x is given by  
给定一个向量（测量样本）x=（x1，…，xn）∈rn，回想一下x的平均值（或平均值）x由

.  
.

We let x − x denote the centered data point  
我们让x-x表示中心数据点。

x − x = (x1 − x,...,xn − x).  
x−x=（x1−x，…，xn−x）。

In order to measure the spread of the xi’s around the mean, we define the sample variance (for short, variance) var(x) (or s2) of the sample x by  
为了测量XI在平均值附近的传播，我们定义了样本X的样本方差（简称为方差）VaR（x）（或S2）。

Example 21.6. If 2), and var(x) =  
例21.6。如果2）和var（x）=

1), and var(  
1）和var（

2.  
2。

There is a reason for using n − 1 instead of n. The above definition makes var(x) an unbiased estimator of the variance of the random variable being sampled. However, we don’t need to worry about this. Curious readers will find an explanation of these peculiar definitions in Epstein [58] (Chapter 14, Section 14.5) or in any decent statistics book.  
使用n−1而不是n是有原因的。上述定义使var（x）成为被采样随机变量方差的无偏估计量。不过，我们不必为此担心。好奇的读者会在爱泼斯坦[58]中（第14章，第14.5节）或任何像样的统计书中找到对这些特殊定义的解释。

Given two vectors x = (x1,...,xn) and y = (y1,...,yn), the sample covariance (for short, covariance) of x and y is given by  
给定两个向量x=（x1，…，xn）和y=（y1，…，yn），x和y的样本协方差（简称协方差）由下式给出：

cov(.  
冠状病毒

Example 21.7. If we take x = (1,3,−1) and y = (0,2,−2), we know from Example 21.6 that x − x = (0,2,−2) and y − y = (−1,0,1). Thus, cov(x,y) = 0(−1)+2(0)+(2 −2)(1) = −1.  
例21.7。如果我们取x=（1,3、-1）和y=（0,2、-2），我们从例21.6中知道x−x=（0,2、-2）和y−y=（-1,0,1）。因此，cov（x，y）=0（−1）+2（0）+（2−2）（1）=1。

The covariance of x and y measures how x and y vary from the mean with respect to each other. Obviously, cov(x,y) = cov(y,x) and cov(x,x) = var(x).  
x和y的协方差度量x和y在平均值上的差异。显然，cov（x，y）=cov（y，x）和cov（x，x）=var（x）。

Note that  
注意

cov(.  
冠状病毒

We say that x and y are uncorrelated iff cov(x,y) = 0.  
我们说x和y是不相关的，iff cov（x，y）=0。

Finally, given an n × d matrix X of n points Xi, for PCA to be meaningful, it will be necessary to translate the origin to the centroid (or center of gravity) µ of the Xi’s, defined by  
最后，给定n点x的n×d矩阵x，对于PCA是有意义的，有必要将原点翻译成Xi的质心（或重心），由

.  
.

Observe that if µ = (µ1,...,µd), then µj is the mean of the vector Cj (the jth column of X).  
观察，如果μ=（μ1，…，μd），则μj是矢量Cj（x的jth列）的平均值。

We let X − µ denote the matrix whose ith row is the centered data point Xi − µ (1 ≤ i ≤ n). Then the sample covariance matrix (for short, covariance matrix) of X is the d × d symmetric matrix  
我们以X为表示矩阵为Iz行为中心的数据点Xi（1）那么x的样本协方差矩阵（简称协方差矩阵）就是d×d对称矩阵。

.  
.

Example 21.8. Let, the 3 × 2 matrix whose columns are the vector x and −  
例21.8。设3×2矩阵，其列为矢量x和−

y of Example 21.6. Then  
例21.6的y。然后

,  
，

and  
和

.  
.

Remark: The factor is irrelevant for our purposes and can be ignored.  
备注：该因素与我们的目的无关，可以忽略不计。

Example 21.9. Here is the matrix X −µ in the case of our bearded mathematicians: since  
例21.9。下面是我们的胡须数学家的矩阵x−μ：因为

µ1 = 1828.4, µ2 = 5.6,  
礹1=1828.4，礹2=5.6，

we get  
我们得到

|  |  |  |
| --- | --- | --- |
| Name 网络错误 | year 网络错误 | length 网络错误 |
| Carl Friedrich Gauss 网络错误 | −51.4 网络错误 | −5.6 网络错误 |
| Camille Jordan 网络错误 | 9.6 网络错误 | 6.4 网络错误 |
| Adrien-Marie Legendre 网络错误 | −76.4 网络错误 | −5.6 网络错误 |
| Bernhard Riemann 网络错误 | −2.4 网络错误 | 9.4 网络错误 |
| David Hilbert 网络错误 | 33.6 网络错误 | −3.6 网络错误 |
| Henri Poincar´e 网络错误 | 25.6 网络错误 | −0.6 网络错误 |
| Emmy Noether 网络错误 | 53.6 网络错误 | −5.6 网络错误 |
| Karl Weierstrass 网络错误 | 13.4 网络错误 | −5.6 网络错误 |
| Eugenio Beltrami 网络错误 | 6.6 网络错误 | −3.6 网络错误 |
| Hermann Schwarz 网络错误 | 14.6 网络错误 | 14.4 网络错误 |

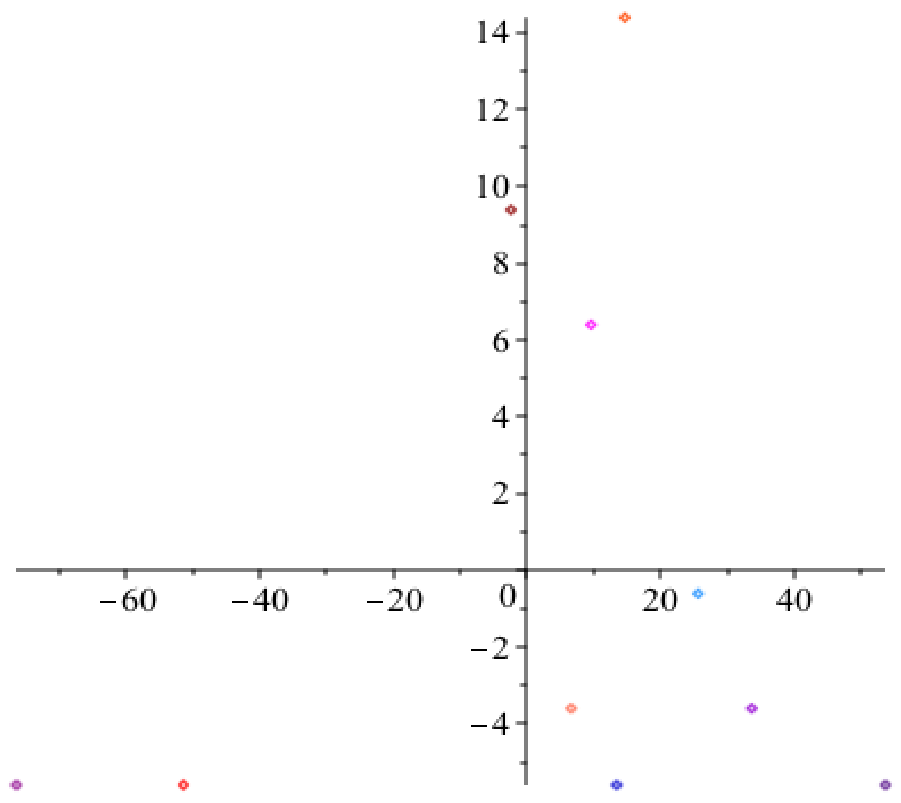
See Figure 21.3.  
见图21.3。

We can think of the vector Cj as representing the features of X in the direction ej (the jth canonical basis vector in Rd, namely ej = (0,...,1,...0), with a 1 in the jth position).  
我们可以把矢量Cj看作是表示x在ej方向上的特征（rd中的jth规范基矢量，即ej=（0，…，1，…，0），其中1在jth位置）。

If v ∈ Rd is a unit vector, we wish to consider the projection of the data points X1,...,Xn onto the line spanned by v. Recall from Euclidean geometry that if x ∈ Rd is any vector and v ∈ Rd is a unit vector, the projection of x onto the line spanned by v is hx,viv.  
如果v∈rd是单位向量，我们希望考虑数据点x1，…，xn在v所跨越的直线上的投影。从欧几里德几何中回忆，如果x∈rd是任何向量，v∈rd是单位向量，x在v所跨越的直线上的投影是hx，viv。

Thus, with respect to the basis v, the projection of x has coordinate hx,vi. If x is represented by a row vector and v by a column vector, then  
因此，对于基V，x的投影具有坐标hx，vi。如果x由行向量表示，v由列向量表示，那么

hx,vi = xv.  
hx，vi=xv。



Gauss

Jordan

Legendre

Riemann

Hilbert

Poincare

Noether

Weierstrass

Beltrami

Schwarz

Figure 21.3: The centered data points of Example 21.9.  
图21.3：示例21.9的中心数据点。

Therefore, the vector Y ∈ Rn consisting of the coordinates of the projections of X1,...,Xn onto the line spanned by v is given by Y = Xv, and this is the linear combination  
因此，由x1，…，xn的投影坐标构成的向量y∈rn在V所跨越的线上，由y=xv给出，这是线性组合。

Xv = v1C1 + ··· + vdCd  
xv=v1c1+·····+vdcd

of the columns of X (with v = (v1,...,vd)).  
x列（v=（v1，…，vd））。

Observe that because µj is the mean of the vector Cj (the jth column of X), we get  
观察到，由于μj是矢量cj（x的jth列）的平均值，我们得到

Y = Xv = v1µ1 + ··· + vdµd,  
y=xv=v1祄1+····+vd祄d，

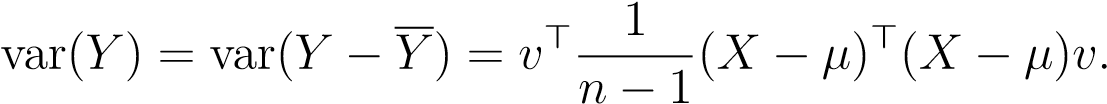
and so the centered point Y − Y is given by  
因此中心点y−y由

Y − Y = v1(C1 − µ1) + ··· + vd(Cd − µd) = (X − µ)v.  
Y−Y=v1（c1−μ1）+·····+v d（cd−μd）=（x−μ）V。

Furthermore, if Y = Xv and Z = Xw, then  
此外，如果y=xv，z=xw，那么

cov(  
冠状病毒

where Σ is the covariance matrix of X. Since Y − Y has zero mean, we have  
其中∑是x的协方差矩阵。由于y−y的均值为零，我们得到



The above suggests that we should move the origin to the centroid µ of the Xi’s and consider the matrix X − µ of the centered data points Xi − µ.  
这意味着我们应该把原点移动到Xi的质心上，并考虑中心数据点Xi的矩阵X。

From now on beware that we denote the columns of X − µ by C1,...,Cd and that Y denotes the centered point is a unit vector.  
从现在开始要注意，我们用c1，…，cd表示x-，y表示中心点是单位向量。

Basic idea of PCA: The principal components of X are uncorrelated projections Y of the data points X1, ..., Xn onto some directions v (where the v’s are unit vectors) such that var(Y ) is maximal.  
主成分分析的基本思想：x的主要成分是数据点x1，…，xn的不相关投影y到一些方向v（其中v是单位向量），使得var（y）最大。

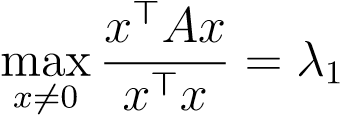
This suggests the following definition:  
这表明了以下定义：

Definition 21.2. Given an n×d matrix X of data points X1,...,Xn, if µ is the centroid of the Xi’s, then a first principal component of X (first PC) is a centered point Y1 = (X−µ)v1, the projection of X1,...,Xn onto a direction v1 such that var(Y1) is maximized, where v1 is a unit vector (recall that Y1 = (X − µ)v1 is a linear combination of the Cj’s, the columns of X − µ).  
定义21.2.给定数据点X1，…，Xn的N×D矩阵X，如果X是XI的质心，那么X（第一PC）的第一主分量是中心点Y1=（xω）V1，X1，…，Xn投影到方向V1，使得VAR（Y1）被最大化，其中V1是单位向量（回忆Y1＝（x）。-μ）v1是Cj的线性组合，x-μ的柱）。

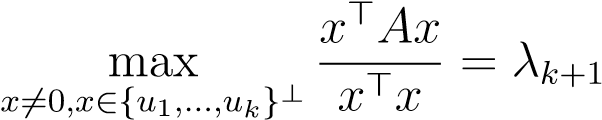
More generally, if Y1,...,Yk are k principal components of X along some unit vectors v1,...,vk, where 1 ≤ k < d, a (k+1)th principal component of X ((k+1)th PC) is a centered point Yk+1 = (X − µ)vk+1, the projection of X1,...,Xn onto some direction vk+1 such that var(Yk+1) is maximized, subject to cov(Yh,Yk+1) = 0 for all h with 1 ≤ h ≤ k, and where vk+1 is a unit vector (recall that Yh = (X − µ)vh is a linear combination of the Cj’s). The vh are called principal directions.  
更一般地说，如果y1，…，yk是x的k主分量，沿着一些单位向量v1，…，vk，其中1≤k<d，x的a（k+1）th主分量（（k+1）th pc）是一个中心点yk+1=（x-μ）vk+1，x1，…，xn在某个方向上的投影vk+1，这样var（yk+1）最大化，subject to cov（yh，yk+1）=0，对于1≤h≤k的所有h，其中vk+1是单位矢量（回想一下，yh=（x−µ）vh是cj的线性组合）。VH被称为主方向。

The following proposition is the key to the main result about PCA. This result was already proven in Proposition 16.23 except that the eigenvalues were listed in increasing order. For the reader’s convenience we prove it again.  
下面的命题是主成分分析主要结果的关键。这一结果已经在命题16.23中得到证明，只是特征值是按递增顺序列出的。为了读者的方便，我们再次证明了这一点。

Proposition 21.10. If A is a symmetric d × d matrix with eigenvalues λ1 ≥ λ2 ≥ ··· ≥ λd and if (u1,...,ud) is any orthonormal basis of eigenvectors of A, where ui is a unit eigenvector associated with λi, then  
提案21.10。如果a是特征值λ1≥λ2≥·····················································



(with the maximum attained for x = u1) and  
（X=U1时达到最大值）和



(with the maximum attained for x = uk+1), where 1 ≤ k ≤ d − 1.  
（X=UK+1时达到最大值），其中1≤K≤D−1。

Proof. First observe that  
证据。首先要注意

,  
，

and similarly,  
同样地，

.  
.

Since A is a symmetric matrix, its eigenvalues are real and it can be diagonalized with respect to an orthonormal basis of eigenvectors, so let (u1,...,ud) be such a basis. If we write  
由于A是一个对称矩阵，其特征值是实的，它可以相对于特征向量的正态基对角化，因此（u1，…，ud）就是这样的基。如果我们写信

,  
，

a simple computation shows that  
简单的计算表明

d x>Ax = Xλix2i .  
d x>ax=xλix2i。

i=1  
i＝1

If x>x = 1, then = 1, and since we assumed that λ1 ≥ λ2 ≥ ··· ≥ λd, we get  
如果x>x=1，则=1，由于我们假设λ1≥λ2≥·································

.  
.

Thus,  
因此，

,  
，

and since this maximum is achieved for e1 = (1,0,...,0), we conclude that  
由于e1（1,0，…，0）达到了这个最大值，我们得出结论

.  
.

Next observe that x ∈ {u1,...,uk}⊥ and x>x = 1 iff x1 = ··· = xk = 0 and  
接下来观察x∈u1，…，uk和x>x=1 iff x1=····=xk=0和

Consequently, for such an x, we have  
因此，对于这样一个x，我们有

.  
.

Thus,  
因此，

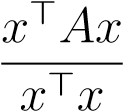
,  
，

and since this maximum is achieved for ek+1 = (0,...,0,1,0,...,0) with a 1 in position k+1, we conclude that  
由于Ek+1的最大值为（0，…，0,1,0，…，0），位置k+1为1，我们得出结论：

,  
，

as claimed.   
如要求。

The quantity  
数量



is known as the Rayleigh ratio or Rayleigh–Ritz ratio (see Section 16.6 ) and Proposition 21.10 is often known as part of the Rayleigh–Ritz theorem.  
被称为瑞利比或瑞利-瑞兹比（见第16.6节），而21.10命题通常被称为瑞利-瑞兹定理的一部分。

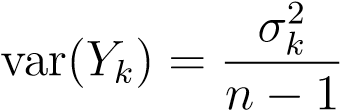
Proposition 21.10 also holds if A is a Hermitian matrix and if we replace x>Ax by x∗Ax and x>x by x∗x. The proof is unchanged, since a Hermitian matrix has real eigenvalues and is diagonalized with respect to an orthonormal basis of eigenvectors (with respect to the Hermitian inner product).  
命题21.10也适用于如果a是厄米特矩阵，并且如果我们用x ax替换x>ax，x>x替换x x，则证明是不变的，因为厄米特矩阵具有实际特征值，并且相对于特征向量的正交基（关于厄米特内积）对角化。

We then have the following fundamental result showing how the SVD of X yields the PCs:  
然后，我们将得到以下基本结果，说明X的SVD如何生成PC：

Theorem 21.11. (SVD yields PCA) Let X be an n × d matrix of data points X1,...,Xn, and let µ be the centroid of the Xi’s. If X − µ = V DU> is an SVD decomposition of X − µ and if the main diagonal of D consists of the singular values σ1 ≥ σ2 ≥ ··· ≥ σd, then the centered points Y1,...,Yd, where  
定理21.11。（SVD产生PCA），X为数据点X1，…，Xn的N×D矩阵，并设为Xi的质心。如果X＝＝V DU>是X×SVD的SVD分解，如果D的主对角线由奇异值α1×2以上的±·ω-δD组成，则中心点Y1，…，YD，WH。埃尔

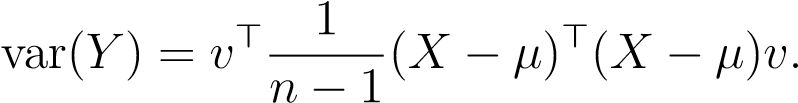
Yk = (X − µ)uk = kth column of V D  
yk=（x−μ）uk=v d的第k列

and uk is the kth column of U, are d principal components of X. Furthermore,  
Uk是u的第k列，是x的d个主要成分。

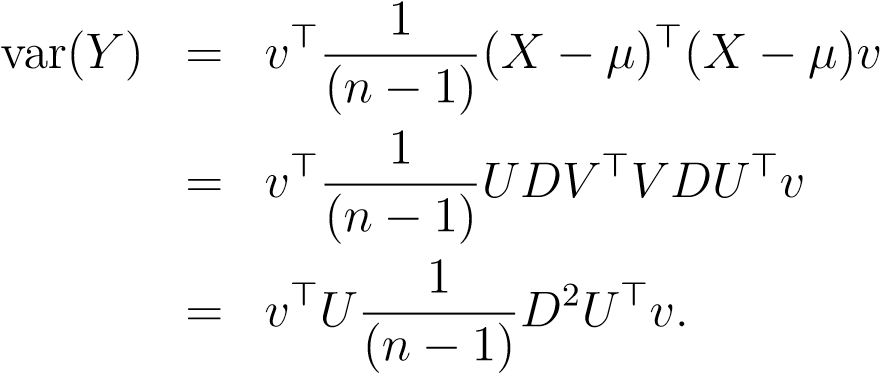


and cov(Yh,Yk) = 0, whenever h =6 k and 1 ≤ k,h ≤ d.  
而cov（yh，yk）=0，当h=6 k且1≤k时，h≤d。

Proof. Recall that for any unit vector v, the centered projection of the points X1,...,Xn onto the line of direction v is Y = (X − µ)v and that the variance of Y is given by  
证据。回想一下，对于任何单位向量v，点x1，…，xn在方向v的直线上的中心投影是y=（x-μ）v，y的方差由下式得出：



Since X − µ = V DU>, we get  
由于x−μ=v du>，我们得到

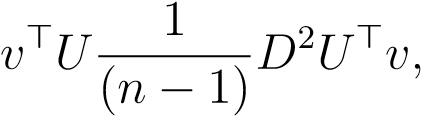


Similarly, if Y = (X − µ)v and Z = (X − µ)w, then the covariance of Y and Z is given by  
同样，如果y=（x−μ）v和z=（x−μ）w，则y和z的协方差由下式得出：

cov(  
冠状病毒

the columns ofObviously, U (n−1U1)Dform an orthonormal basis of unit eigenvectors.2U> is a symmetric matrix whose eigenvalues are , and  
实际上，u（n−1u1）数据列构成单位特征向量的正态基。2u>是一个对称矩阵，其特征值为，和

We proceed by induction on k. For the base case, k = 1, maximizing var(Y ) is equivalent to maximizing  
我们对k进行归纳。对于基本情况，k=1，最大化var（y）等于最大化



where v is a unit vector. By Proposition 21.10, the maximum of the above quantity is the largest eigenvalue of, namely, and it is achieved for u1, the first columnn of U. Now we get  
其中v是单位向量。由命题21.10可知，上述数量的最大值是最大特征值，即，对于u的第一列u1，我们得到

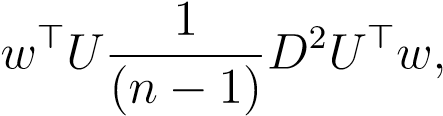
Y1 = (X − µ)u1 = V DU>u1,  
y1=（x−µ）u1=v du>u1，

and since the columns of U form an orthonormal basis, U>u1 = e1 = (1,0,...,0), and so Y1 is indeed the first column of V D.  
由于u的列构成正交基，u>u1=e1=（1,0，…，0），所以y1确实是v d的第一列。

By the induction hypothesis, the centered points Y1,...,Yk, where Yh = (X − µ)uh and u1,...,uk are the first k columns of U, are k principal components of X. Because  
根据诱导假设，中心点y1，…，yk，其中yh=（x−µ）uh和u1，…，u k是u的前k列，是x的k主要成分。因为

cov(  
冠状病毒

where Y = (X − µ)v and Z = (X − µ)w, the condition cov(Yh,Z) = 0 for h = 1,...,k is equivalent to the fact that w belongs to the orthogonal complement of the subspace spanned by {u1,...,uk}, and maximizing var(Z) subject to cov(Yh,Z) = 0 for h = 1,...,k is equivalent to maximizing  
式中，y=（x−μ）v和z=（x−μ）w，条件cov（y h，z）=0，对于h=1，…，k等于w属于由u1，…，uk所跨越的子空间的正交补集，并且服从cov（yh，z）=0，对于h=1，…，k等于最大化。



where w is a unit vector orthogonal to the subspace spanned by {u1,...,uk}. By Proposition  
其中w是一个与子空间正交的单位向量，其范围为u1，…，uk。按命题

21.10, the maximum of the above quantity is the (k+1)th eigenvalue of, namely  
21.10，上述数量的最大值为（k+1）的第（k+1）个特征值，即

, and it is achieved for uk+1, the (k + 1)th columnn of U. Now we get  
它是为英国+1，美国的第（k+1）列而实现的。

Yk+1 = (X − µ)uk+1 = V DU>uk+1,  
YK+1=（x−µ）UK+1=V du>UK+1，

and since the columns of U form an orthonormal basis, U>uk+1 = ek+1, and Yk+1 is indeed the (k + 1)th column of V D, which completes the proof of the induction step.   
由于u列构成正交基，u>u k+1=ek+1，yk+1确实是v d的（k+1）第（k+1）列，完成了诱导步骤的证明。

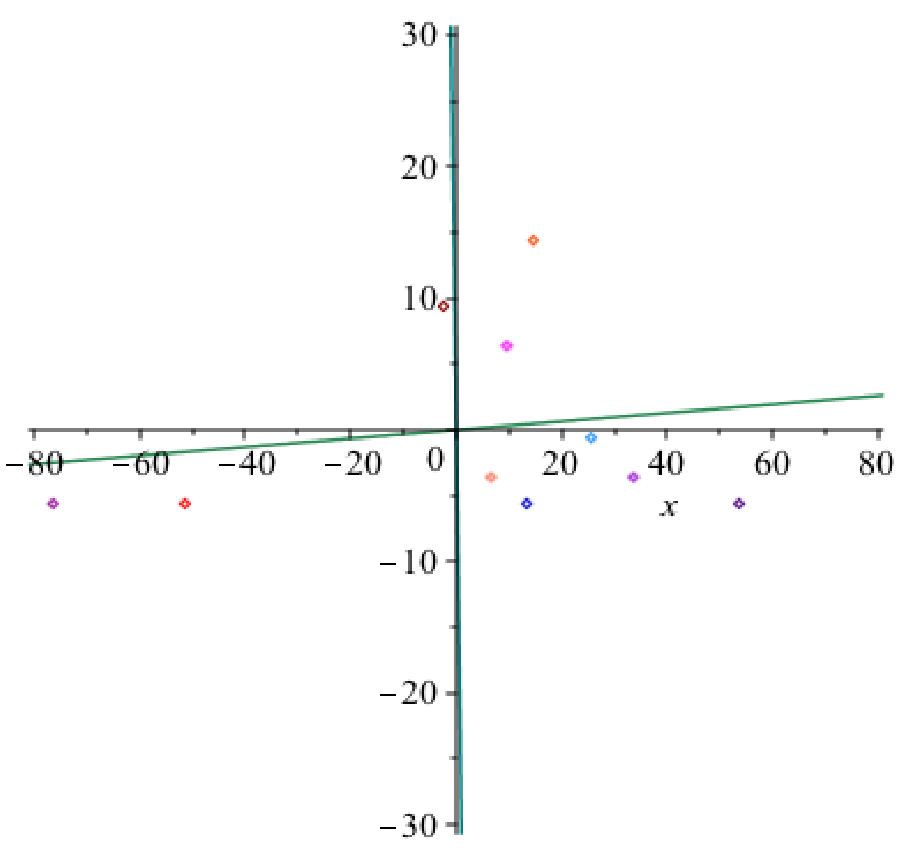
The d columns u1,...,ud of U are usually called the principal directions of X − µ (and X). We note that not only do we have cov(Yh,Yk) = 0 whenever h =6 k, but the directions u1,...,ud along which the data are projected are mutually orthogonal.  
u的d列u1，…，ud通常称为x−礹（和x）的主方向。我们注意到，不仅当h=6K时，cov（yh，yk）=0，而且数据投影的方向u1，…，ud是相互正交的。

Example 21.10. For the centered data set of our bearded mathematicians (Example 21.9) we have X − µ = V ΣU>, where Σ has two nonzero singular values, σ1 = 116.9803,σ2 = 21.7812, and with  
例21.10。对于胡须数学家的中心数据集（例21.9），我们有x−µ=v∑u>，其中∑有两个非零奇异值，σ1=116.9803，σ2=21.7812，以及

,  
，

so the principal directions are u1 = (0.9995,0.0325) and u2 = (0.0325,−0.9995). Observe that u1 is almost the direction of the x-axis, and u2 is almost the opposite direction of the y-axis. We also find that the projections Y1 and Y2 along the principal directions are  
所以主方向是U1=（0.9995,0.0325）和U2=（0.0325，−0.9995）。观察u1几乎是x轴的方向，u2几乎是y轴的相反方向。我们还发现沿着主方向的投影y1和y2是

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| −51.5550 网络错误   9.8031 −76.5417 网络错误   网络错误   −2.0929 网络错误   网络错误   33.4651 V D =  25.5669 网络错误   网络错误   网络错误   53.3894 网络错误   网络错误   13.2107 网络错误   网络错误   6.4794 网络错误  15.0607 网络错误 | 3.9249  网络错误  −6.0843  网络错误  3.1116  网络错误  −9.4731  网络错误  4.6912 , 网络错误  1.4325  7.3408  网络错误  6.0330  网络错误  3.8128  −13.9174 网络错误 | with 网络错误 | −51.4000 网络错误   9.6000 −76.4000 网络错误   网络错误   −2.4000 网络错误   网络错误   33.6000 网络错误  Xµ =  25.6000 网络错误   网络错误   网络错误   53.6000 网络错误   网络错误   13.4000 网络错误   网络错误   6.6000 网络错误  14.6000 网络错误 | −5.6000 网络错误  6.4000  网络错误  −5.6000 网络错误  9.4000  −3.6000. −0.6000 网络错误  −5.6000 网络错误  −5.6000 网络错误  −3.6000 网络错误  14.4000 网络错误 |
| See Figures 21.4, 21.5, and 21.6. 网络错误 | |



u

1

u

2

Gauss

Legendre

Riemann

Jordan

Schwarz

Noether

Weierstrass

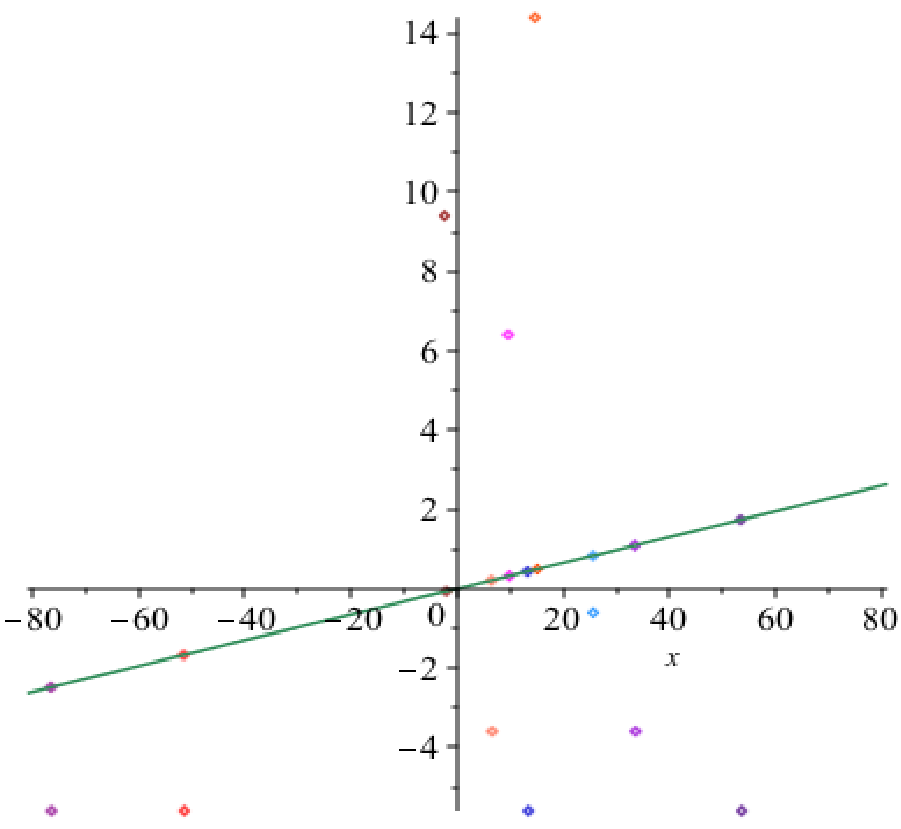
Hilbert

Poincaire

Beltrami

Figure 21.4: The centered data points of Example 21.9 and the two principal directions of Example 21.10.  
图21.4：实施例21.9的中心数据点和实施例21.10的两个主要方向。

We know from our study of SVD that are the eigenvalues of the symmetric positive semidefinite matrix (X − µ)>(X − µ) and that u1,...,ud are corresponding eigenvectors. Numerically, it is preferable to use SVD on X −µ rather than to compute explicitly (Xµ) and then diagonalize it. Indeed, the explicit computation of A>A from a matrix can be numerically quite unstable, and good SVD algorithms avoid computing A>A explicitly.  
我们从对SVD的研究中知道，对称半正定矩阵（x−礹）>（x−礹）的特征值和u1，…，ud是相应的特征向量。在数值上，最好在x−μ上使用SVD，而不是显式计算（xμ），然后对其进行对角线化。事实上，从矩阵中显式计算a>a可能在数值上相当不稳定，并且良好的SVD算法避免显式计算a>a。



Gauss

Jordan

Schwarz

Poincaire

Legendre

Beltrami

Riemann

Hilbert

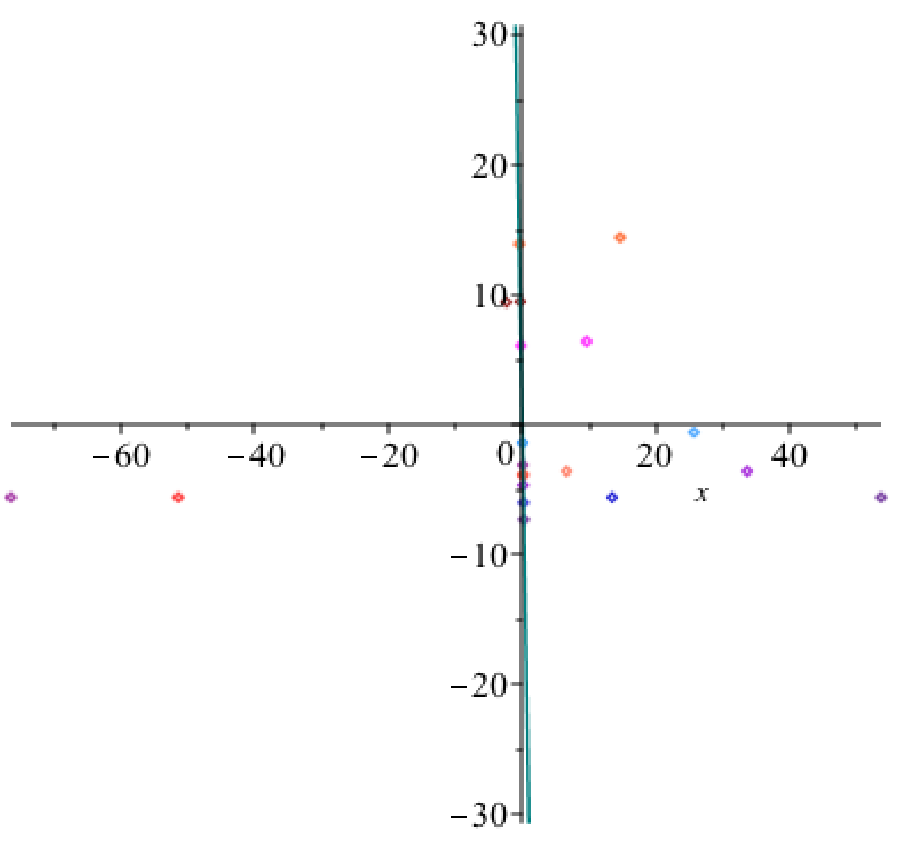
Noether

Weierstrass

u

1

Figure 21.5: The first principal components of Example 21.10, i.e. the projection of the centered data points onto the u1 line.  
图21.5：实施例21.10的第一个主要组成部分，即中心数据点在U1线上的投影。



Legendre

Gauss

Riemann

Jordan

Schwarz

Beltrami

Weierstrass

Poincare

Hilbert

Noether

u

2

Figure 21.6: The second principal components of Example 21.10, i.e. the projection of the centered data points onto the u2 line.  
图21.6：实施例21.10的第二个主要组成部分，即中心数据点在U2线上的投影。

In general, since an SVD of X is not unique, the principal directions u1,...,ud are not unique. This can happen when a data set has some rotational symmetries, and in such a case, PCA is not a very good method for analyzing the data set.  
一般来说，由于x的svd不是唯一的，所以主方向u1，…，ud不是唯一的。当数据集具有某些旋转对称性时，就会发生这种情况，在这种情况下，PCA不是一种很好的数据集分析方法。

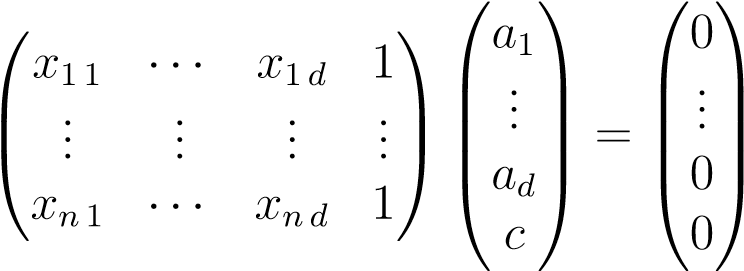
## 21.5 Best Affine Approximation 21.5最佳仿射近似

A problem very close to PCA (and based on least squares) is to best approximate a data set of n points X1,...,Xn, with Xi ∈ Rd, by a p-dimensional affine subspace A of Rd, with 1 ≤ p ≤ d − 1 (the terminology rank d − p is also used).  
一个非常接近PCA的问题（和基于最小二乘法）是最好地逼近一个N点X1，…，Xn的数据集，用X-RD，用RD的p维仿射子空间A，用1个±p＝D 1（术语等级D P）也使用。

First consider p = d−1. Then A = A1 is an affine hyperplane (in Rd), and it is given by an equation of the form  
首先考虑P=D-1。那么a=a1是仿射超平面（在rd中），它由形式方程给出。

a1x1 + ··· + adxd + c = 0.  
A1x1+·····+ADxd+c=0.

By best approximation, we mean that (a1,...,ad,c) solves the homogeneous linear system  
通过最佳逼近，我们的意思是（a1，…，ad，c）解齐次线性系统。



in the least squares sense, subject to the condition that a = (a1,...,ad) is a unit vector, that is, a>a = 1, where Xi = (xi1,··· ,xid). If we form the symmetric matrix  
在最小二乘意义上，服从A=（A1，…，AD）是单位向量的条件，即A＞A＝1，其中X=（XI1，FAY·，XID）。如果我们形成对称矩阵

x11 ··· > x11 ··· x1d 1  
x11········································

x1d 1  
X1D 1

.  
.

 .. ... ... ...  ... ... ... ...  
………………………………………………

xn1 ··· xnd 1 xn1 ··· xnd 1  
xn1···xnd 1 xn1···xnd 1

involved in the normal equations, we see that the bottom row (and last column) of that matrix is  
在正规方程中，我们看到矩阵的最下面一行（和最后一列）是

nµ1 ··· nµd n,  
n礹1···n礹d n，

where times the mean of the column Cj of X.  
其中乘以x的cj列的平均值。

Therefore, if (a1,...,ad,c) is a least squares solution, that is, a solution of the normal equations, we must have  
因此，如果（a1，…，ad，c）是一个最小二乘解，也就是说，一个正态方程的解，我们必须

nµ1a1 + ··· + nµdad + nc = 0,  
n礹a1+·····+n礹dad+nc=0，

that is, a1µ1 + ··· + adµd + c = 0,  
也就是说，A1礹1+·····+AD礹d+C=0，

which means that the hyperplane A1 must pass through the centroid µ of the data points X1,...,Xn. Then we can rewrite the original system with respect to the centered data Xi − µ, find that the variable c drops out, get the system (X − µ)a = 0,  
这意味着超平面A1必须通过数据点x1，…，xn的质心。然后，我们可以重写原始系统相对于中心数据Xi，发现变量C退出，得到系统（x＝\*）A＝0，

### 21.5. BEST AFFINE APPROXIMATION 21.5。最佳仿射近似

where a = (a1,...,ad).  
其中a=（a1，…，ad）。

Thus, we are looking for a unit vector a solving (X − µ)a = 0 in the least squares sense, that is, some a such that a>a = 1 minimizing  
因此，我们在寻找一个单位向量a，在最小二乘意义上求解（x−μ）a=0，也就是说，一些a，使得a>a=1最小化

a>(X − µ)>(X − µ)a.  
A>（X−礹）>（X−礹）A.

Compute some SVD V DU> of X −µ, where the main diagonal of D consists of the singular values σ1 ≥ σ2 ≥ ··· ≥ σd of X − µ arranged in descending order. Then  
计算x−μ的一些svd v du>值，其中d的主对角线由x−μ的奇异值σ1≥σ2≥·······························然后

a>(X − µ)>(X − µ)a = a>UD2U>a,  
a>（x−礹）>（x−礹）a=a>ud2u>a，

where D2 = diag() is a diagonal matrix, so pick a to be the last column in U  
其中d2=diag（）是一个对角矩阵，所以选择a作为u中的最后一列

(corresponding to the smallest eigenvalue σd2 of (X − µ)>(X − µ)). This is a solution to our best fit problem.  
（对应于（x−μ）>（x−μ）的最小特征值σd2）。这是我们最适合的问题的解决方案。

Therefore, if Ud−1 is the linear hyperplane defined by a, that is,  
因此，如果ud−1是由a定义的线性超平面，也就是说，

Ud−1 = {u ∈ Rd | hu,ai = 0},  
ud−1=u∈rd hu，ai=0，

where a is the last column in U for some SVD V DU> of X − µ, we have shown that the affine hyperplane A1 = µ + Ud−1 is a best approximation of the data set X1,...,Xn in the least squares sense.  
其中a是x−μ的某些svd v du>的u中的最后一列，我们已经证明仿射超平面a1=μ+ud−1是最小平方意义上的数据集x1，…，xn的最佳近似值。

It is easy to show that this hyperplane A1 = µ + Ud−1 minimizes the sum of the square distances of each Xi to its orthogonal projection onto A1. Also, since Ud−1 is the orthogonal complement of a, the last column of U, we see that Ud−1 is spanned by the first d−1 columns of U, that is, the first d − 1 principal directions of X − µ.  
很容易证明，超平面A1=ω+UD 1最小化了每个XI的平方距离与A1上的正交投影的平方之和。此外，由于u d−1是u的最后一列a的正交补码，我们发现ud−1由u的第一个d−1列（即x−µ的第一个d−1主方向）构成。

All this can be generalized to a best (d−k)-dimensional affine subspace Ak approximating X1,...,Xn in the least squares sense (1 ≤ k ≤ d−1). Such an affine subspace Ak is cut out by k independent hyperplanes Hi (with 1 ≤ i ≤ k), each given by some equation  
所有这些都可以推广到一个最佳（d−k）维仿射子空间ak，在最小二乘意义上近似于x1，…，xn（1≤k≤d−1）。这样的仿射子空间ak由k独立超平面hi（1≤i≤k）切出，每个超平面由一些方程给出。

ai1x1 + ··· + aidxd + ci = 0.  
ai1x1+·····+aidxd+ci=0.

If we write ai = (ai1,··· ,aid), to say that the Hi are independent means that a1,...,ak are linearly independent. In fact, we may assume that a1,...,ak form an orthonormal system.  
如果我们写ai=（ai1，···，aid），表示hi是独立的，意味着a1，…，ak是线性独立的。事实上，我们可以假设a1，…，ak形成一个正态系统。

Then finding a best (d − k)-dimensional affine subspace Ak amounts to solving the homogeneous linear system  
然后找到一个最佳的（d−k）维仿射子空间AK等于解齐次线性系统。

,  
，

in the least squares sense, subject to the conditions a>i aj = δij, for all i,j with 1 ≤ i,j ≤ k, where the matrix of the system is a block diagonal matrix consisting of k diagonal blocks (X,1), where 1 denotes the column vector (1,...,1) ∈ Rn.  
在最小二乘意义上，在a>i a j=δij的条件下，对于所有i，j，1≤i，j≤k，其中系统矩阵是由k个对角块（x，1）组成的块对角矩阵，其中1表示列向量（1，…，1）∈rn。

Again it is easy to see that each hyperplane Hi must pass through the centroid µ of X1,...,Xn, and by switching to the centered data Xi − µ we get the system      
很容易看出，每一个超平面Hi都必须通过x1，…，xn的质心，通过切换到中心数据Xi，我们得到了系统。

X − µ 0 ··· 0 a1 0  
X−0···0 A1 0

 ... ... ... ...  ...  = ...,  
……………………=…，

  
γ

0 0 ··· X − µ ak 0  
0 0·····X····AK 0

with a>i aj = δij for all i,j with 1 ≤ i,j ≤ k.  
a>i a j=δij，对于所有i，j，1≤i，j≤k。

If V DU> = X−µ is an SVD decomposition, it is easy to see that a least squares solution of this system is given by the last k columns of U, assuming that the main diagonal of D consists of the singular values σ1 ≥ σ2 ≥ ··· ≥ σd of X−µ arranged in descending order. But now the (d−k)-dimensional subspace Ud−k cut out by the hyperplanes defined by a1,...,ak is simply the orthogonal complement of (a1,...,ak), which is the subspace spanned by the first d − k columns of U.  
如果v d u>=x−µ是一个SVD分解，很容易看出这个系统的最小二乘解是由u的最后k列给出的，假设d的主对角线由x−µ的奇异值σ1≥σ2≥···································但是现在，由a1，…，ak定义的超平面切出的（d−k）维子空间ud−k只是（a1，…，ak）的正交补码，它是由u的第一个d−k列所跨越的子空间。

So the best (d−k)-dimensional affine subpsace Ak approximating X1,...,Xn in the least squares sense is  
因此，在最小二乘意义上，最好的（d−k）维仿射子簇ak近似于x1，…，xn是

Ak = µ + Ud−k,  
ak=μ+ud−k，

where Ud−k is the linear subspace spanned by the first d−k principal directions of X−µ, that is, the first d−k columns of U. Consequently, we get the following interesting interpretation of PCA (actually, principal directions):  
其中，u d−k是由x−μ的第一个d−k主方向（即u的第一个d−k列）所跨越的线性子空间。因此，我们得到了以下有趣的PCA解释（实际上，主方向）：

Theorem 21.12. Let X be an n × d matrix of data points X1,...,Xn, and let µ be the centroid of the Xi’s. If X − µ = V DU> is an SVD decomposition of X − µ and if the main diagonal of D consists of the singular values σ1 ≥ σ2 ≥ ··· ≥ σd, then a best (d − k)dimensional affine approximation Ak of X1,...,Xn in the least squares sense is given by  
定理21.12。设X是数据点X1、…、Xn的N×D矩阵，并设为Xi的质心。如果X＝＝V DU>是X×SVD的SVD分解，如果D的主对角线是由奇异值α1×2以上的ω-ω-ωd，则是一个最好的（d×k）维仿射逼近A。k的x1，…，xn在最小二乘意义上由下式给出

Ak = µ + Ud−k,  
ak=μ+ud−k，

where Ud−k is the linear subspace spanned by the first d − k columns of U, the first d − k principal directions of X − µ (1 ≤ k ≤ d − 1).  
其中，u d−k是由u的第一个d−k列跨越的线性子空间，x−（1≤k≤d−1）的第一个d−k主方向。

Example 21.11. Going back to Example 21.10, a best 1-dimensional affine approximation A1 is the affine line passing through (µ1,µ2) = (1824.4,5.6) of direction u1 = (0.9995,0.0325).  
例21.11。回到实施例21.10，最好的一维仿射近似值a1是穿过U1=（0.9995,0.0325）方向（μ1，μ2）=（1824.4,5.6）的仿射线。

There are many applications of PCA to data compression, dimension reduction, and pattern analysis. The basic idea is that in many cases, given a data set X1,...,Xn, with Xi ∈ Rd, only a “small” subset of m < d of the features is needed to describe the data set accurately.  
PCA在数据压缩、降维和模式分析中有许多应用。其基本思想是，在许多情况下，给定数据集X1，…，Xn，Xi，RD RD，只有一个“小”子集的特征的MD D需要准确地描述数据集。

### 21.6. SUMMARY 21.6。总结

If u1,...,ud are the principal directions of X −µ, then the first m projections of the data (the first m principal components, i.e., the first m columns of V D) onto the first m principal directions represent the data without much loss of information. Thus, instead of using the original data points X1,...,Xn, with Xi ∈ Rd, we can use their projections onto the first m principal directions Y1,...,Ym, where Yi ∈ Rm and m < d, obtaining a compressed version of the original data set.  
如果U1，…，ud是x−祄的主方向，那么数据的第一个m投影（第一个m主分量，即v d的第一个m列）在第一个m主方向上表示数据，而不会丢失太多信息。因此，代替使用原始数据点X1，…，Xn，用Xi×RD，我们可以将它们的投影应用到第一M主方向Y1，…，YM，其中Yi-Rm和M< D，获得原始数据集的压缩版本。

For example, PCA is used in computer vision for face recognition. Sirovitch and Kirby (1987) seem to be the first to have had the idea of using PCA to compress facial images. They introduced the term eigenpicture to refer to the principal directions, ui. However, an explicit face recognition algorithm was given only later by Turk and Pentland (1991). They renamed eigenpictures as eigenfaces.  
例如，PCA用于计算机视觉中的人脸识别。Sirovitch和Kirby（1987）似乎是第一个想到使用PCA压缩面部图像的人。他们引入了“本征图”这个术语来指代主方向，即用户界面。然而，只有在Turk和Pentland（1991）之后才给出了一种明确的人脸识别算法。他们把本征图片改名为本征面。

For details on the topic of eigenfaces, see Forsyth and Ponce [65] (Chapter 22, Section 22.3.2), where you will also find exact references to Turk and Pentland’s papers.  
有关Eigenfaces主题的详细信息，请参阅Forsyth和Ponce[65]（第22章，第22.3.2节），在这里您还可以找到Turk和Pentland论文的确切参考。

Another interesting application of PCA is to the recognition of handwritten digits. Such an application is described in Hastie, Tibshirani, and Friedman, [87] (Chapter 14, Section  
PCA的另一个有趣的应用是手写数字的识别。这种应用在黑斯提、提比西拉尼和弗里德曼[87]中有描述（第14章，第

14.5.1).  
14.5.1条）。

## 21.6 Summary 21.6总结

The main concepts and results of this chapter are listed below:  
本章的主要概念和结果如下：

* Least squares problems.  
  最小二乘问题。
* Existence of a least squares solution of smallest norm (Theorem 21.1).  
  最小范数最小二乘解的存在性（定理21.1）。
* The pseudo-inverse A+ of a matrix A. • The least squares solution of smallest norm is given by the pseudo-inverse (Theorem  
  矩阵A的伪逆A+。•最小范数的最小二乘解由伪逆（定理）给出。

21.2)  
21.2）

* Projection properties of the pseudo-inverse.  
  伪逆的投影属性。
* The pseudo-inverse of a normal matrix.  
  正态矩阵的伪逆矩阵。
* The Penrose characterization of the pseudo-inverse.  
  伪逆的彭罗斯特征。
* Data compression and SVD.  
  数据压缩和SVD。
* Best approximation of rank < r of a matrix.  
  矩阵秩小于r的最佳近似。
* Principal component analysis.  
  主成分分析。
* Review of basic statistical concepts: mean, variance, covariance, covariance matrix.  
  回顾基本统计概念：均值、方差、协方差、协方差矩阵。
* Centered data, centroid.  
  中心数据，质心。
* The principal components (PCA).  
  主要成分（PCA）。
* The Rayleigh–Ritz theorem (Theorem 21.10).  
  瑞利-里兹定理（定理21.10）。
* The main theorem: SVD yields PCA (Theorem 21.11).  
  主定理：SVD产生PCA（定理21.11）。
* Best affine approximation.  
  最佳仿射近似。
* SVD yields a best affine approximation (Theorem 21.12).  
  SVD产生最佳仿射近似（定理21.12）。
* Face recognition, eigenfaces.  
  人脸识别，特征面。

## 21.7 Problems 21.7问题

Problem 21.1. Consider the overdetermined system in the single variable x:  
问题21.1。考虑单变量x中的超定系统：

a1x = b1,...,amx = bm.  
a1x=b1，…，amx=bm。

Prove that the least squares solution of smallest norm is given by  
证明了最小范数的最小二乘解由

.  
.

Problem 21.2. Let X be an m × n real matrix. For any strictly positive constant K > 0, the matrix X>X +KIn is invertible. Prove that the limit of the matrix (X>X +KIn)−1X> when K goes to zero is equal to the pseudo-inverse X+ of X.  
问题21.2。设x为m×n实矩阵。对于任何严格正常数k>0，矩阵x>x+kin是可逆的。证明当k为零时，矩阵（x>x+kin）−1X>的极限等于x的伪逆x+。

Problem 21.3. Use Matlab to find the pseudo-inverse of the 8 × 6 matrix  
问题21.3。用matlab求8×6矩阵的伪逆

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 64 网络错误   9 网络错误  17 网络错误   网络错误  40 A = 32 网络错误   网络错误   网络错误  41 网络错误   网络错误  49 网络错误  8 网络错误 | 2 网络错误  55 网络错误  47 网络错误  26 网络错误  34 网络错误  23 网络错误  15 网络错误  58 网络错误 | 3 54 网络错误  46 网络错误  27 网络错误  35 网络错误  22 网络错误  14 网络错误  59 网络错误 | 61 网络错误  12 网络错误  20 网络错误  37 网络错误  29 网络错误  44 网络错误  52 网络错误  5 网络错误 | 60 网络错误  13 网络错误  21 网络错误  36 网络错误  28 网络错误  45 网络错误  53 网络错误  4 网络错误 | 6  网络错误  51 网络错误  43 网络错误  30 38. 网络错误  19 网络错误  11 网络错误  62 网络错误 |

Observe that the sums of the columns are all equal to to 256. Let b be the vector of dimension 6 whose coordinates are all equal to 256. Find the solution x+ of the system Ax = b.  
观察各列的总和均等于256。设b为坐标均等于256的维度6的向量。找到系统的解决方案x+，ax=b。

Problem 21.4. The purpose of this problem is to show that Proposition 21.9 (the Eckart– Young theorem) also holds for the Frobenius norm. This problem is adapted from Strang  
问题21.4。这个问题的目的是证明21.9命题（Eckart-Young定理）也适用于Frobenius规范。这个问题是根据Strang改编的

[166], Section I.9.  
[166]，第I.9节。

### 21.7. PROBLEMS 21.7。问题

Suppose the m×n matrix B of rank at most k minimizes kA − BkF . Start with an SVD of B,  
假设秩至多k的m×n矩阵b使ka−bkf最小化。从B的SVD开始，

,  
，

where D is a diagonal k × k matrix. We can write  
其中d是对角线k×k矩阵。我们可以写信

,  
，

where L is strictly lower triangular in the first k rows, E is diagonal, and R is strictly upper triangular, and let  
其中，在前k行中，l是严格的下三角形，e是对角的，r是严格的上三角形，并

,  
，

which clearly has rank  
很明显有排名

1. Prove that  
   证明这一点

.  
.

Since kA − BkF is minimal, show that L = R = F = 0.  
因为kA−bkf是最小的，所以表明l=r=f=0。

Similarly, show that G = 0.  
同样，显示g=0。

1. We have  
   我们有

,  
，

where E is diagonal, so deduce that  
其中e是对角线，所以推断

1. D = diag(σ1,...,σk).  
   d=diag（σ1，…，σk）。
2. The singular values of H must be the smallest n − k singular values of A.  
   h的奇异值必须是a的最小n-k奇异值。
3. The minimum of kA − BkF must be.  
   kA−bkf的最小值必须为。

Problem 21.5. Prove that the closest rank 1 approximation (in k k2) of the matrix  
问题21.5。证明矩阵的最近秩1近似（k k2）

is  
是

.  
.

1 matrixShow that the Eckart–Young theorem fails for the operator normB such that kA − Bk∞ < kA − A1k∞. k k∞ by finding a rank Problem 21.6. Find a closest rank 1 approximation (in k k2) for the matrices  
1矩阵：通过发现秩问题21.6，Eckart–Young定理对算符normb失败，从而使k a−bk∞<ka−a1k∞.k k k∞。求矩阵的最近秩1近似值（k k2）

.  
.

Problem 21.7. Find a closest rank 1 approximation (in k k2) for the matrix  
问题21.7。求矩阵的最近秩1近似值（k k2）

.  
.

Problem 21.8. Let S be a real symmetric positive definite matrix and let S = UΣU> be a diagonalization of S. Prove that the closest rank 1 matrix (in the L2-norm) to, where u1 is the first column of U.  
问题21.8。设为实对称正定矩阵，设s=u∑u>为s的对角化，证明最接近的秩1矩阵（在l2范数中），其中u1是u的第一列。

Chapter 22  
第二十二章

# Computing Eigenvalues and Eigenvectors 计算特征值和特征向量

After the problem of solving a linear system, the problem of computing the eigenvalues and the eigenvectors of a real or complex matrix is one of most important problems of numerical linear algebra. Several methods exist, among which we mention Jacobi, Givens–Householder, divide-and-conquer, QR iteration, and Rayleigh–Ritz; see Demmel [49], Trefethen and Bau [171], Meyer [122], Serre [151], Golub and Van Loan [80], and Ciarlet [41]. Typically, better performing methods exist for special kinds of matrices, such as symmetric matrices.  
在求解线性系统问题之后，计算实矩阵或复矩阵的特征值和特征向量的问题是数值线性代数中最重要的问题之一。存在几种方法，其中我们提到Jacobi、Givens——户主、分而治之、QR迭代和Rayleigh——Ritz；见Demmel[49]、Trefetten和Bau[171]、Meyer[122]、Serre[151]、Golub和van Loan[80]和Ciarlet[41]。通常，对于特殊类型的矩阵（如对称矩阵），存在性能更好的方法。

In theory, given an n×n complex matrix A, if we could compute a Schur form A = UTU∗, where T is upper triangular and U is unitary, we would obtain the eigenvalues of A, since they are the diagonal entries in T. However, this would require finding the roots of a polynomial, but methods for doing this are known to be numerically very unstable, so this is not a practical method.  
在理论上，给定一个n×n复矩阵a，如果我们可以计算一个Schur形式a=u t u，其中t是上三角，u是单位的，我们就可以得到a的特征值，因为它们是t中的对角项。然而，这需要求多项式的根，但要求出方法这是众所周知的数值非常不稳定，所以这不是一个实际的方法。

A common paradigm is to construct a sequence (Pk) of matrices such that Ak = Pk−1APk converges, in some sense, to a matrix whose eigenvalues are easily determined. For example, Ak = Pk−1APk could become upper triangular in the limit. Furthermore, Pk is typically a product of “nice” matrices, for example, orthogonal matrices.  
一个常见的范例是构造一个矩阵序列（pk），使得ak=pk−1apk在某种意义上收敛到一个特征值容易确定的矩阵。例如，ak=pk−1apk可能在极限处变成上三角形。此外，pk通常是“nice”矩阵的乘积，例如正交矩阵。

For general matrices, that is, matrices that are not symmetric, the QR iteration algorithm, due to Rutishauser, Francis, and Kublanovskaya in the early 1960s, is one of the most efficient algorithms for computing eigenvalues. A fascinating account of the history of the QR algorithm is given in Golub and Uhlig [79]. The QR algorithm constructs a sequence of matrices (Ak), where Ak+1 is obtained from Ak by performing a QR-decomposition Ak = QkRk of Ak and then setting Ak+1 = RkQk, the result of swapping Qk and Rk. It is immediately verified that Ak+1 = Q∗kAkQk, so Ak and Ak+1 have the same eigenvalues, which are the eigenvalues of A.  
对于一般的矩阵，即非对称矩阵，20世纪60年代初由于Rutishauser、Francis和Kublanovskaya的影响，QR迭代算法是计算特征值最有效的算法之一。Golub和Uhlig[79]中给出了QR算法历史的精彩描述。QR算法构造了一个矩阵序列（AK），其中AK+1是通过对AK执行QR分解AK=QKRK，然后设置AK+1=RKQK，交换QK和RK的结果从AK获得的。立即证实，ak+1=q kakqk，因此ak和ak+1具有相同的特征值，即a的特征值。

The basic version of this algorithm runs into difficulties with matrices that have several eigenvalues with the same modulus (it may loop or not “converge” to an upper triangular matrix). There are ways of dealing with some of these problems, but for ease of exposition,  
该算法的基本版本在具有多个具有相同模的特征值的矩阵中遇到困难（它可能循环或不“收敛”到上三角矩阵）。有一些方法可以解决这些问题，但为了便于解释，

663  
六百六十三

we first present a simplified version of the QR algorithm which we call basic QR algorithm. We prove a convergence theorem for the basic QR algorithm, under the rather restrictive hypothesis that the input matrix A is diagonalizable and that its eigenvalues are nonzero and have distinct moduli. The proof shows that the part of Ak strictly below the diagonal converges to zero and that the diagonal entries of Ak converge to the eigenvalues of A.  
我们首先提出了一个简化版的二维码算法，我们称之为基本二维码算法。在输入矩阵A可对角化且特征值不为零且具有明显模性的限制性假设下，证明了基本QR算法的收敛定理。证明了严格低于对角的AK部分收敛到零，AK的对角项收敛到A的特征值。

Since the convergence of the QR method depends crucially only on the fact that the part of Ak below the diagonal goes to zero, it would be highly desirable if we could replace A by a similar matrix U∗AU easily computable from A and having lots of zero strictly below the diagonal. It turns out that there is a way to construct a matrix H = U∗AU which is almost triangular, except that it may have an extra nonzero diagonal below the main diagonal. Such matrices called, Hessenberg matrices, are discussed in Section 22.2. An n×n diagonalizable Hessenberg matrix H having the property that hi+1i = 06 for i = 1,...,n − 1 (such a matrix is called unreduced) has the nice property that its eigenvalues are all distinct. Since every Hessenberg matrix is a block diagonal matrix of unreduced Hessenberg blocks, it suffices to compute the eigenvalues of unreduced Hessenberg matrices. There is a special case of particular interest: symmetric (or Hermitian) positive definite tridiagonal matrices. Such matrices must have real positive distinct eigenvalues, so the QR algorithm converges to a diagonal matrix.  
由于qr方法的收敛性主要取决于一个事实，即对角线下的ak部分变为零，因此，如果我们可以用一个类似的矩阵u au替换a，该矩阵u au很容易从a计算出来，并且在对角线下有大量的零，这将是非常可取的。事实证明，有一种方法可以构造一个几乎是三角形的矩阵h=u au，除了它在主对角线下面可能有一个额外的非零对角线。在第22.2节中讨论了这种称为Hessenberg矩阵的矩阵。一个n×n的可对角化Hessenberg矩阵h，其性质为i=1，…，n−1（这种矩阵称为无约矩阵）具有其特征值都不同的优良性质。由于每一个海森堡矩阵都是一个由海森堡块组成的块对角矩阵，所以计算海森堡矩阵的特征值就足够了。有一个特别有趣的例子：对称（或厄米特）正定三对角矩阵。这样的矩阵必须具有实正的特征值，因此QR算法收敛到一个对角矩阵。

In Section 22.3, we consider techniques for making the basic QR method practical and more efficient. The first step is to convert the original input matrix A to a similar matrix H in Hessenberg form, and to apply the QR algorithm to H (actually, to the unreduced blocks of H). The second and crucial ingredient to speed up convergence is to add shifts.  
在第22.3节中，我们考虑了使基本QR方法更实用、更有效的技术。第一步是将原始输入矩阵A转换为类似的Hessenberg形式的矩阵H，并将QR算法应用于H（实际上，应用于H的未减少块）。加速收敛的第二个关键因素是增加移位。

A shift is the following step: pick some σk, hopefully close to some eigenvalue of A (in general, λn), QR-factor Ak − σkI as  
移动是以下步骤：选择一些σk，希望接近a（一般来说，λn）的特征值，qr因子ak-σki as

Ak − σkI = QkRk,  
AK−σki=qkrk，

and then form  
然后形成

Ak+1 = RkQk + σkI.  
AK+1=RKQK+σki。

It is easy to see that we still have Ak+1 = Q∗kAkQk. A judicious choice of σk can speed up convergence considerably. If H is real and has pairs of complex conjugate eigenvalues, we can perform a double shift, and it can be arranged that we work in real arithmetic.  
很容易看出我们仍然有ak+1=q kakqk。明智地选择σk可以大大加快收敛速度。如果h是实的，并且有一对复共轭特征值，我们可以执行双移位，并且可以安排我们在实算术中工作。

The last step for improving efficiency is to compute Ak+1 = Q∗kAkQk without even performing a QR-factorization of Ak −σkI. This can be done when Ak is unreduced Hessenberg. Such a method is called QR iteration with implicit shifts. There is also a version of QR iteration with implicit double shifts.  
提高效率的最后一步是计算ak+1=q kakqk，甚至不执行ak−σki的qr因子分解。这可以在AK是非公爵海森堡时完成。这种方法称为隐式移位的QR迭代。还有一个带有隐式双移位的QR迭代版本。

If the dimension of the matrix A is very large, we can find approximations of some of the eigenvalues of A by using a truncated version of the reduction to Hessenberg form due to Arnoldi in general and to Lanczos in the symmetric (or Hermitian) tridiagonal case. Arnoldi iteration is discussed in Section 22.4. If A is an m × m matrix, for much smaller than m) the idea is to generate the n × n Hessenberg submatrix Hn of the full Hessenberg matrix H (such that A = UHU∗) consisting of its first n rows and n columns; the matrix Un consisting of the first n columns of U is also produced. The Rayleigh–Ritz method consists in computing the eigenvalues of Hn using the QR- method with shifts. These eigenvalues, called Ritz values, are approximations of the eigenvalues of A. Typically, extreme eigenvalues are found first.  
如果矩阵A的维数非常大，我们可以通过使用截断形式的约简来找到A的一些特征值的近似值，这种约简形式通常是由于阿诺迪和兰佐斯在对称（或厄米提亚）三对角情况下的约简。第22.4节讨论了Arnoldi迭代。如果a是m×m矩阵，对于比m小得多的矩阵，其思想是生成完整的Hessenberg矩阵h（这样a=u h u）的n×n Hessenberg子矩阵hn，该子矩阵由其前n行和n列组成；也生成不由u的前n列组成的矩阵。瑞利-瑞兹方法是利用位移的QR-方法计算hn的特征值。这些特征值称为Ritz值，是A特征值的近似值。通常首先找到极端特征值。

Arnoldi iteration can also be viewed as a way of computing an orthonormal basis of a Krylov subspace, namely the subspace Kn(A,b) spanned by (b,Ab,...,Anb). We can also use Arnoldi iteration to find an approximate solution of a linear equation Ax = b by minimizing kb − Axnk2 for all xn is the Krylov space Kn(A,b). This method named GMRES is discussed in Section 22.5.  
Arnoldi迭代也可以看作是计算krylov子空间的正态基的一种方法，即（b，ab，…，anb）所跨越的子空间kn（a，b）。我们也可以使用Arnoldi迭代，通过最小化所有xn的kb−axnk2，找到线性方程ax=b的近似解，即krylov空间kn（a，b）。第22.5节讨论了名为gmres的方法。

The special case where H is a symmetric (or Hermitian) tridiagonal matrix is discussed in Section 22.6. In this case, Arnoldi’s algorithm becomes Lanczos’ algorithm. It is much more efficient than Arnoldi iteration.  
在第22.6节中讨论了H是对称（或厄米特）三对角矩阵的特殊情况。在这种情况下，Arnoldi的算法变成了Lanczos的算法。它比Arnoldi迭代更有效。

We close this chapter by discussing two classical methods for computing a single eigenvector and a single eigenvalue: power iteration and inverse (power) iteration; see Section  
我们通过讨论计算单个特征向量和单个特征值的两种经典方法来结束本章：功率迭代和逆（功率）迭代；参见第节

22.7.  
22.7。

## 22.1 The Basic QR Algorithm 22.1基本QR算法

Let A be an n × n matrix which is assumed to be diagonalizable and invertible. The basic QR algorithm makes use of two very simple steps. Starting with A1 = A, we construct sequences of matrices (Ak), (Qk) (Rk) and (Pk) as follows:  
假设a是一个n×n矩阵，它假定是对角化的和可逆的。基本的QR算法使用两个非常简单的步骤。从a1=a开始，我们构造矩阵（ak）、（qk）（rk）和（pk）的序列，如下所示：

Factor A1 = Q1R1  
系数a1=q1r1

Set A2 = R1Q1  
设置a2=r1q1

Factor A2 = Q2R2  
系数a2=q2r2

Set A3 = R2Q2  
设置a3=r2q2

...  
…

Factor Ak = QkRk  
系数ak=qkrk

Set Ak+1 = RkQk  
设置AK+1=RKQK

...  
…

Thus, Ak+1 is obtained from a QR-factorization Ak = QkRk of Ak by swapping Qk and  
因此，通过交换qk和

Rk. Define Pk by  
RK。定义pk的依据

Pk = Q1Q2 ···Qk.  
pk=q1q2···qk。

Since Ak = QkRk, we have , and since Ak+1 = RkQk, we obtain  
因为ak=qkrk，我们有，并且因为ak+1=rkqk，我们得到

. (∗1)  
.（1）

An obvious induction shows that  
一个明显的归纳显示

,  
，

that is  
那就是

. (∗2)  
.（2）

Therefore, Ak+1 and A are similar, so they have the same eigenvalues.  
因此，AK+1和A是相似的，所以它们具有相同的特征值。

The basic QR iteration method consists in computing the sequence of matrices Ak, and in the ideal situation, to expect that Ak “converges” to an upper triangular matrix, more precisely that the part of Ak below the main diagonal goes to zero, and the diagonal entries converge to the eigenvalues of A.  
基本的QR迭代方法包括计算矩阵的序列AK，在理想情况下，期望AK“收敛”到上三角矩阵，更精确地说，主对角线下的AK部分变为零，对角线条目收敛到特征值。a.

This ideal situation is only achieved in rather special cases. For one thing, if A is unitary (or orthogonal in the real case), since in the QR decomposition we have R = I, we get A2 = IQ = Q = A1, so the method does not make any progress. Also, if A is a real matrix, since the Ak are also real, if A has complex eigenvalues, then the part of Ak below the main diagonal can’t go to zero. Generally, the method runs into troubles whenever A has distinct eigenvalues with the same modulus.  
这种理想情况只有在相当特殊的情况下才能实现。首先，如果a是一元的（或在实际情况下是正交的），因为在q r分解中我们有r=i，我们得到a2=i q=q=a1，所以这个方法没有任何进展。另外，如果a是一个实矩阵，因为ak也是实的，如果a有复杂的特征值，那么主对角线下面的ak部分就不能归零。一般来说，当A的特征值不同且模相同时，该方法就会遇到麻烦。

The convergence of the sequence (Ak) is only known under some fairly restrictive hypotheses. Even under such hypotheses, this is not really genuine convergence. Indeed, it can be shown that the part of Ak below the main diagonal goes to zero, and the diagonal entries converge to the eigenvalues of A, but the part of Ak above the diagonal may not converge. However, for the purpose of finding the eigenvalues of A, this does not matter.  
序列的收敛性（ak）只有在一些相当严格的假设下才能知道。即使在这样的假设下，这也不是真正的趋同。事实上，可以证明主对角线下的ak部分为零，对角线条目收敛到a的特征值，但对角线上的ak部分可能不收敛。然而，为了求A的特征值，这并不重要。

The following convergence result is proven in Ciarlet [41] (Chapter 6, Theorem 6.3.10 and Serre [151] (Chapter 13, Theorem 13.2). It is rarely applicable in practice, except for symmetric (or Hermitian) positive definite matrices, as we will see shortly.  
以下收敛结果在Ciarlet[41]中得到证明（第6章，定理6.3.10和Serre[151]（第13章，定理13.2）。它在实践中很少适用，除了对称（或厄米特）正定矩阵，我们稍后将看到。

Theorem 22.1. Suppose the (complex) n×n matrix A is invertible, diagonalizable, and that its eigenvalues λ1,...,λn have different moduli, so that  
定理22.1。假设（复数）n×n矩阵a是可逆的、可对角化的，且其特征值λ1，…，λn具有不同的模，因此

|λ1| > |λ2| > ··· > |λn| > 0.  
|λ1>λ2>···>λn>0.

If A = PΛP −1, where Λ = diag(λ1,...,λn), and if P −1 has an LU-factorization, then the strictly lower-triangular part of Ak converges to zero, and the diagonal of Ak converges to Λ.  
如果a=p∧p−1，其中∧=diag（λ1，…，λn），并且如果p−1具有Lu因式分解，那么AK的严格下三角部分收敛到零，AK的对角线收敛到∧。

Proof. We reproduce the proof in Ciarlet [41] (Chapter 6, Theorem 6.3.10). The strategy is to study the asymptotic behavior of the matrices Pk = Q1Q2 ···Qk. For this, it turns out that we need to consider the powers Ak.  
证据。我们在Ciarlet[41]中复制了证明（第6章，定理6.3.10）。研究矩阵pk=q1q2···qk的渐近行为。为此，我们需要考虑权力AK。

Step 1. Let Rk = Rk ···R2R1. We claim that  
第1步。设Rk=Rk···r2r1。我们声称

Ak = (Q1Q2 ···Qk)(Rk ···R2R1) = PkRk. (∗3)  
ak=（q1q2···qk）（rk···r2r1）=pkrk。（3）

We proceed by induction. The base case k = 1 is trivial. For the induction step, from  
我们采用归纳法。基本情况k=1无关紧要。对于感应步骤，从

(∗2), we have  
（2），我们有

PkAk+1 = APk.  
pkak+1=apk。

Since Ak+1 = RkQk = Qk+1Rk+1, we have  
由于AK+1=RKQK=QK+1RK+1，我们有

Pk+1Rk+1 = PkQk+1Rk+1Rk = PkAk+1Rk = APkRk = AAk = Ak+1  
pk+1rk+1=pkqk+1rk+1rk=pk ak+1rk=apkrk=aak=ak+1

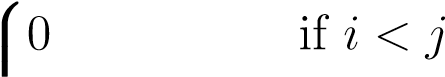
establishing the induction step.  
建立诱导步骤。

Step 2. We will express the matrix Pk as Pk = QQekDk, in terms of a diagonal matrix Dk with unit entries, with Q and Qek unitary.  
第2步。我们将矩阵pk表示为pk=q qek dk，用带单位项的对角矩阵dk表示，q和qek为一元。

Let P = QR, a QR-factorization of P (with R an upper triangular matrix with positive diagonal entries), and P −1 = LU, an LU-factorization of P −1. Since A = PΛP −1, we have  
设p=qr，p的qr因子分解（r为上三角矩阵，带正对角项），p−1=lu，p−1的lu因子分解。既然a=p∧p−1，我们有

Ak = PΛkP −1 = QRΛkLU = QR(ΛkLΛ−k)ΛkU. (∗4)  
ak=p∧kp−1=qr∧klu=qr（∧kl∧k）∧ku.（4）

Here, Λ−k is the diagonal matrix with entries λ−i k. The reason for introducing the matrix ΛkLΛ−k is that its asymptotic behavior is easy to determine. Indeed, we have  
这里，∧−k是条目为λ−i k的对角矩阵。引入矩阵∧kl∧−k的原因是其渐近行为易于确定。事实上，我们有



kLΛ−k)ij = 1 k if i = j  
kl∧−k）i j=1 k，如果i=j

(Λ  
（b）

The hypothesis that |λ1| > |λ2| > ··· > |λn| > 0 implies that  
假设λ1>λ2>····>λn>0意味着

lim ΛkLΛ−k = I. (†) k7→∞  
lim∧kl∧−k=i.（†）k7→∞

Note that it is to obtain this limit that we made the hypothesis on the moduli of the eigenvalues. We can write  
注意，为了得到这个极限，我们假设了特征值的模。我们可以写信

ΛkLΛ−k = I + Fk, with lim Fk = 0,  
∧kl∧−k=i+fk，其中lim fk=0，

k7→∞  
K7→∞

and consequently, since R(ΛkLΛ−k) = R(I + Fk) = R + RFkR−1R = (I + RFkR−1)R, we have  
因此，由于r（∧kl∧k）=r（i+fk）=r+rfkr−1r=（i+rfkr−1）r，我们得出

R(ΛkLΛ−k) = (I + RFkR−1)R. (∗5)  
R（∧kl∧k）=（I+Rfkr−1）R.（5）

By Proposition 8.11(1), since limk7→∞ Fk = 0, and thus limk7→∞ RFkR−1 = 0, the matrices I + RFkR−1 are invertible for k large enough. Consequently for k large enough, we have a QR-factorization  
根据命题8.11（1），由于limk7→∞fk=0，因此limk7→∞rfkr−1=0，矩阵i+rfkr−1对于k足够大是可逆的。因此，对于足够大的k，我们有一个qr因子分解。

I + RFkR−1 = QekRek, (∗6)  
i+rfkr−1=qekrek，（6）

with (Rek)ii > 0 for i = 1,...,n. Since the matrices Qek are unitary, we have = 1, so the sequence (Qek) is bounded. It follows that it has a convergent subsequence (Qe`) that converges to some matrix Qe, which is also unitary. Since  
当（rek）i i>0时，i=1，…，n。因为矩阵qek是一元的，所以我们有=1，所以序列（qek）是有界的。因此，它有一个收敛子序列（qe`），它收敛到一个矩阵qe，这个矩阵qe也是一元的。自从

Re` = (Qe`)∗(I + RF`R−1),  
re`=（qe`）（i+rf`r−1），

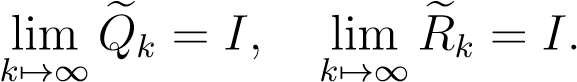
we deduce that the subsequence (Re`) also converges to some matrix Re, which is also upper triangular with positive diagonal entries. By passing to the limit (using the subsequences), we get Re = (Qe)∗, that is,  
我们推导出子序列（re`）也收敛于某个矩阵re，该矩阵也是上三角形，具有正对角项。通过传递到极限（使用子序列），我们得到re=（qe），也就是说，

I = QeR.e  
i=qer.e

By the uniqueness of a QR-decomposition (when the diagonal entries of R are positive), we get  
通过qr分解的唯一性（当r的对角项为正时），我们得到

Qe = Re = I.  
qe=re=i。

Since the above reasoning applies to any subsequences of (Qek) and (Rek), by the uniqueness of limits, we conclude that the “full” sequences (Qek) and (Rek) converge:  
由于上述推理适用于（qek）和（rek）的任何子序列，根据极限的唯一性，我们得出结论，“全”序列（qek）和（rek）收敛：



Since by (∗4),  
从（4）开始，

Ak = QR(ΛkLΛ−k)ΛkU,  
ak=qr（∧kl∧k）∧ku，

by (∗5),  
通过（5）

R(ΛkLΛ−k) = (I + RFkR−1)R,  
R（∧kl∧k）=（I+Rfkr−1）R，

and by (∗6)  
以及（6）

I + RFkR−1 = QekRek,  
i+rfkr−1=qekrek，

we proved that  
我们证明了

Ak = (QQek)(RekRΛkU). (∗7)  
ak=（qqek）（rekr∧ku）。（7）

Observe that QQek is a unitary matrix, and RekRΛkU is an upper triangular matrix, as a product of upper triangular matrices. However, some entries in Λ may be negative, so we can’t claim that has positive diagonal entries. Nevertheless, we have another QR-decomposition of Ak,  
观察到qqek是一个单位矩阵，rekr∧ku是一个上三角矩阵，作为上三角矩阵的乘积。然而，∧中的一些条目可能是负数，所以我们不能声称有正对角线条目。不过，我们还有另一个AK的QR分解，

Ak = (QQek)(RekRΛkU) = PkRk.  
ak=（qqek）（rekr∧ku）=pkrk.

It is easy to prove that there is diagonal matrix Dk with |(Dk)ii| = 1 for i = 1,...,n, such that  
很容易证明存在对角矩阵dk，其中i=1，…，n的（dk）ii=1，这样

. (∗8)  
.（8）

The existence of Dk is consequence of the following fact: If an invertible matrix B has two QR factorizations B = Q1R1 = Q2R2, then there is a diagonal matrix D with unit entries such that Q2 = DQ1.  
dk的存在是以下事实的结果：如果可逆矩阵b有两个qr因式分解b=q1r1=q2r2，那么就有一个对角矩阵d，其单位项为q2=dq1。

The expression for Pk in (∗8) is that which we were seeking.  
（8）中pk的表达式是我们正在寻找的表达式。

Step 3. Asymptotic behavior of the matrices Ak+1 = Pk∗APk.  
第3步。矩阵ak+1=pk apk的渐近行为。

Since A = PΛP −1 = QRΛR−1Q−1 and by (∗8), Pk = QQekDk, we get  
由于a=p∧p−1=qr∧r−1q−1和by（8），pk=qqekdk，我们得到

. (∗9)  
.（9）

Since limk7→∞ Qek = I, we deduce that  
由于limk7→∞qek=i，我们推断

|  |  |  |  |
| --- | --- | --- | --- |
|  | 2 网络错误  0 网络错误 | ···... 网络错误  ··· 网络错误 | , 网络错误 |

 λ1 ∗ ···  
λ1···

0 λ  
0兆

an upper triangular matrix with the eigenvalues of A on the diagonal. Since R is upper triangular, the order of the eigenvalues is preserved. If we let  
在对角线上具有a特征值的上三角矩阵。由于R是上三角形，所以特征值的阶数保持不变。如果我们让

Dk = (Qek)∗RΛR−1Qek, (∗10)  
dk=（qek）r∧r−1qek，（10）

then by (∗9) we have Ak+1 = Dk∗DkDk, and since the matrices Dk are diagonal matrices, we have  
然后通过（9），我们得到了ak+1=dk dkdk，由于dk矩阵是对角矩阵，我们得到了

(Ak+1)jj = (Dk∗DkDk)ij = (Dk)ii(Dk)jj(Dk)ij,  
（ak+1）jj=（dk dkdk）ij=（dk）ii（dk）jj（dk）ij，

which implies that  
这意味着

(Ak+1)ii = (Dk)ii, i = 1,...,n, (∗11)  
（ak+1）i i=（dk）ii，i=1，…，n，（11）

since |(Dk)ii| = 1 for i = 1,...,n. Since limk→∞7 Dk = RΛR−1, we conclude that the strictly lower-triangular part of Ak+1 converges to zero, and the diagonal of Ak+1 converges to Λ.   
由于（dk）i i=1，对于i=1，…，n.由于limk→∞7 dk=r∧r−1，我们得出的结论是，ak+1的严格下三角部分收敛到零，而ak+1的对角线收敛到∧。

Observe that if the matrix A is real, then the hypothesis that the eigenvalues have distinct moduli implies that the eigenvalues are all real and simple.  
如果矩阵A是实的，那么特征值具有不同模的假设意味着特征值都是实的和简单的。

The following Matlab program implements the basic QR-method using the function qrv4 from Section 11.8.  
下面的matlab程序使用第11.8节中的函数qrv4实现基本的qr方法。

function T = qreigen(A,m)  
函数t=qreigen（a，m）

T = A; for k = 1:m  
t=a；对于k=1:m

[Q R] = qrv4(T);  
[q r]=qrv4（t）；

T = R\*Q;  
t=r\*q；

end end  
结束

Example 22.1. If we run the function qreigen with 100 iterations on the 8×8 symmetric  
例22.1。如果我们在8×8对称上运行100次迭代的函数qreigen

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| matrix 网络错误 |  |  |  |  |  |  |  |  |
|  |  | 1 网络错误  4 网络错误  1 网络错误  0 网络错误  0 网络错误  0 网络错误  0 网络错误  0 网络错误 | 0 网络错误  1 网络错误  4 网络错误  1 网络错误  0 网络错误  0 网络错误  0 网络错误  0 网络错误 | 0 网络错误  0 网络错误  1 网络错误  4 网络错误  1 网络错误  0 网络错误  0 网络错误  0 网络错误 | 0 网络错误  0 网络错误  0 网络错误  1 网络错误  4 网络错误  1 网络错误  0 网络错误  0 网络错误 | 0 网络错误  0 网络错误  0 网络错误  0 网络错误  1 网络错误  4 网络错误  1 网络错误  0 网络错误 | 0 网络错误  0 网络错误  0 网络错误  0 网络错误  0 网络错误  1 网络错误  4 网络错误  1 网络错误 | 0 网络错误  0 网络错误  0 网络错误  0 0, 网络错误  0 网络错误  1 4 网络错误 |

we find the matrix  
我们找到矩阵

.  
.

The diagonal entries match the eigenvalues found by running the Matlab function eig(A).  
对角线条目与运行matlab函数eig（a）得到的特征值相匹配。

If several eigenvalues have the same modulus, then the proof breaks down, we can no longer claim (†), namely that  
如果几个特征值具有相同的模，那么证明就失效了，我们不能再声称（†），即

lim ΛkLΛ−k = I. k7→∞  
lim∧kl∧−k=i.k7→∞

If we assume that P −1 has a suitable “block LU-factorization,” it can be shown that the matrices Ak+1 converge to a block upper-triangular matrix, where each block corresponds to eigenvalues having the same modulus. For example, if A is a 9 × 9 matrix with eigenvalues λi such that |λ1| = |λ2| = |λ3| > |λ4| > |λ5| = |λ6| = |λ7| = |λ8| = |λ9|, then Ak converges to a block diagonal matrix (with three blocks, a 3 × 3 block, a 1 × 1 block, and a 5 × 5 block)  
如果我们假设P−1有一个合适的“块Lu因子分解”，可以证明矩阵AK+1收敛到块上三角矩阵，其中每个块对应具有相同模的特征值。例如，如果a是一个特征值为λi的9×9矩阵，使得 =124\ \\124;λ1 \\124; \124\ \ λi的特征值为λi的9×9矩阵为\\ \\\\\\\\\\\124; \块）

|  |  |  |  |
| --- | --- | --- | --- |
| of the form 网络错误 |  |  |  |
| ? 网络错误  ? ? 网络错误   网络错误  0 网络错误   网络错误  0 网络错误   网络错误   网络错误  0 网络错误   网络错误  0 网络错误   网络错误  0 网络错误  0 网络错误 | ? 网络错误  ? 网络错误  ? 网络错误  0 网络错误  0 网络错误  0 网络错误  0 网络错误  0 网络错误  0 网络错误 | ? 网络错误  ? 网络错误  ? 网络错误  0 网络错误  0 网络错误  0 网络错误  0 网络错误  0 网络错误  0 网络错误 | ∗ ∗ ∗ ∗ ∗ ∗ 网络错误  ∗ ∗ ∗ ∗ ∗ ∗  网络错误  ∗ ∗ ∗ ∗ ∗ ∗ 网络错误  ? ∗ ∗ ∗ ∗ ∗ 网络错误  0 ? ? ? ? ?. 0 ? ? ? ? ? 网络错误  0 ? ? ? ? ? 网络错误  0 ? ? ? ? ? 0 ? ? ? ? ? 网络错误 |

See Ciarlet [41] (Chapter 6 Section 6.3) for more details.  
更多详情请参见CIARLET[41]（第6章第6.3节）。

Under the conditions of Theorem 22.1, in particular, if A is a symmetric (or Hermitian) positive definite matrix, the eigenvectors of A can be approximated. However, when A is not a symmetric matrix, since the upper triangular part of Ak does not necessarily converge, one has to be cautious that a rigorous justification is lacking.  
在定理22.1的条件下，特别是，如果a是对称（或厄米特）正定矩阵，则a的特征向量可以近似。然而，当a不是对称矩阵时，由于ak的上三角部分不一定收敛，因此必须注意缺乏严格的理由。

Suppose we apply the QR algorithm to a matrix A satisfying the hypotheses of Theorem Theorem 22.1. For k large enough, is nearly upper triangular and the diagonal entries of Ak+1 are all distinct, so we can consider that they are the eigenvalues of Ak+1, and thus of A. To avoid too many subscripts, write T for the upper triangular matrix  
假设我们将QR算法应用于满足定理22.1假设的矩阵A。当k足够大时，是近上三角的，且ak+1的对角项都是不同的，因此我们可以认为它们是ak+1的特征值，因此是a的特征值。为了避免下标太多，请为上三角矩阵写t。

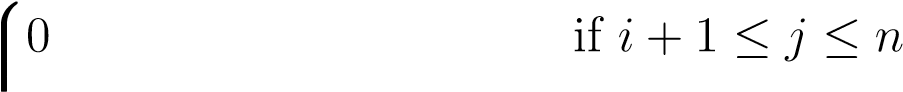
obtained by settting the entries of the part of Ak+1 below the diagonal to 0. Then we can find the corresponding eigenvectors by solving the linear system  
通过将对角线下的AK+1部分的条目设置为0获得。然后通过求解线性系统，求出相应的特征向量。

Tv = tiiv,  
tv=tiiv，

and since T is upper triangular, this can be done by bottom-up elimination. We leave it as an exercise to show that the following vectors ) are eigenvectors:  
因为T是上三角形，所以这可以通过自下而上的消除来实现。我们把它作为一个练习来证明以下向量）是特征向量：

v1 = e1,  
v1=e1，

and if i = 2,...,n, then  
如果i=2，…，n，那么



.  
.

Then the vectors (Pkv1,...,Pkvn) are a basis of (approximate) eigenvectors for A. In the special case where T is a diagonal matrix, then vi = ei for i = 1,...,n and the columns of Pk are an orthonormal basis of (approximate) eigenvectors for A.  
然后向量（pkv1，…，pkvn）是a的特征向量（近似）的基础，在t是对角矩阵的特殊情况下，i=1，…，n的vi=ei和pk的列是a的特征向量（近似）的正交基。

If A is a real matrix whose eigenvalues are not all real, then there is some complex pair of eigenvalues λ + iµ (with µ = 0)6 , and the QR-algorithm cannot converge to a matrix whose strictly lower-triangular part is zero. There is a way to deal with this situation using upper Hessenberg matrices which will be discussed in the next section.  
如果a是一个实矩阵，其特征值不都是实的，则存在一对复杂的特征值λ+i礹（具有礹=0）6，qr算法不能收敛到严格下三角部分为零的矩阵。有一种方法可以使用上赫森堡矩阵来处理这种情况，这将在下一节中讨论。

Since the convergence of the QR method depends crucially only on the fact that the part of Ak below the diagonal goes to zero, it would be highly desirable if we could replace A by a similar matrix U∗AU easily computable from A having lots of zero strictly below the diagonal. We can’t expect U∗AU to be a diagonal matrix (since this would mean that A was easily diagonalized), but it turns out that there is a way to construct a matrix H = U∗AU which is almost triangular, except that it may have an extra nonzero diagonal below the main diagonal. Such matrices called Hessenberg matrices are discussed in the next section.  
由于qr方法的收敛性主要取决于一个事实，即对角线下的ak部分变为零，因此，如果我们可以用一个类似的矩阵u au替换a，则很有必要从严格位于对角线下的具有大量零的矩阵u au进行计算。我们不能期望u au是一个对角矩阵（因为这意味着a很容易对角化），但事实证明有一种方法可以构造一个几乎是三角形的矩阵h=u au，除了它在主对角线下面可能有一个额外的非零对角线。下一节将讨论这种称为Hessenberg矩阵的矩阵。

## 22.2 Hessenberg Matrices 22.2 Hessenberg矩阵

Definition 22.1. An n × n matrix (real or complex) H is an (upper) Hessenberg matrix if it is almost triangular, except that it may have an extra nonzero diagonal below the main diagonal. Technically, hjk = 0 for all (j,k) such that j − k ≥ 2.  
定义22.1.n×n矩阵（实数或复数）h是（上）Hessenberg矩阵，如果它几乎是三角形的，除了在主对角线下面可能有一个额外的非零对角线。从技术上讲，所有（j，k）的hjk=0，因此j−k≥2。

The 5 × 5 matrix below is an example of a Hessenberg matrix.  
下面的5×5矩阵是Hessenberg矩阵的一个例子。

 ∗ ∗ ∗ ∗ ∗  
 ∗ ∗ ∗ ∗ ∗

h21 ∗ ∗ ∗ ∗  
H21

H =  0 h32 ∗ ∗ ∗.  
H=0 H32。

   
 

 0 0 h43 ∗ ∗  
0 0 h43\_\_

0 0 0 h54 ∗  
0 0 0 h54\_

The following result can be shown.  
可以显示以下结果。

Theorem 22.2. Every n × n complex or real matrix A is similar to an upper Hessenberg matrix H, that is, A = UHU∗ for some unitary matrix U. Furthermore, H can be constructed as a product of Householder matrices (the definition is the same as in Section 12.1, except that W is a complex vector, and that the inner product is the Hermitian inner product on Cn). If A is a real matrix, then H is an orthogonal matrix (and H is a real matrix).  
定理22.2.每一个n×n复数或实数矩阵a都类似于上赫森堡矩阵h，也就是说，对于某些单位矩阵u，a=uhu。此外，h可以被构造为户主矩阵的乘积（定义与第12.1节中的定义相同，只是w是一个复数向量，并且innER产品是中国大陆的赫敏内产品。如果a是实矩阵，那么h是正交矩阵（h是实矩阵）。

Theorem 22.2 and algorithms for converting a matrix to Hessenberg form are discussed in Trefethen and Bau [171] (Lecture 26), Demmel [49] (Section 4.4.6, in the real case), Serre [151] (Theorem 13.1), and Meyer [122] (Example 5.7.4, in the real case). The proof of correctness is not difficult and will be the object of a homework problem.  
定理22.2和将矩阵转换为Hessenberg形式的算法在Trefethen和Bau[171]（第26讲）、Demmel[49]（第4.4.6节，在实际情况下）、Serre[151]（定理13.1）和Meyer[122]（在实际情况下，示例5.7.4）中进行了讨论。正确性的证明并不难，而且将成为家庭作业问题的对象。

The following functions written in Matlab implement a function to compute a Hessenberg form of a matrix.  
以下用matlab编写的函数实现了一个计算矩阵Hessenberg形式的函数。

The function house constructs the normalized vector u defining the Householder reflection that zeros all but the first entries in a vector x.  
函数屋构造标准化向量u，定义了户主反射，它将向量x中除第一个条目之外的所有条目归零。

function [uu, u] = house(x) tol = 2\*10^(-15); % tolerance uu = x; p = size(x,1);  
函数[u u，u]=house（x）tol=2\*10^（-15）；%公差uu=x；p=size（x，1）；

% computes l^1-norm of x(2:p,1) n1 = sum(abs(x(2:p,1))); if n1 <= tol  
%计算x的l^1-范数（2:p，1）n1=和（abs（x（2:p，1））；如果n1<=tol

u = zeros(p,1); uu = u;  
u=零（p，1）；uu=u；

else  
其他的

l = sqrt(x’\*x); % l^2 norm of x uu(1) = x(1) + signe(x(1))\*l; u = uu/sqrt(uu’\*uu); end end  
l=sqrt（x'\*x）；%l^2 x u u（1）=x（1）+signe（x（1））\*l；u=uu/sqrt（uu'\*uu）；结束

The function signe(z) returms −1 if z < 0, else +1.  
如果z<0，则函数signe（z）返回−1，否则返回+1。

The function buildhouse builds a Householder reflection from a vector uu.  
buildhouse函数从向量UU构建一个户主反射。

function P = buildhouse(v,i)  
功能P=建筑房屋（V，I）

% This function builds a Householder reflection  
%这个功能建立了一个户主的反映

% [I 0 ] % [0 PP]  
%[I 0]%[0页]

% from a Householder reflection  
%从户主的反映

% PP = I - 2uu\*uu’  
%pp=i-2uu\*uu'

% where uu = v(i:n)  
%式中uu=v（i:n）

% If uu = 0 then P - I  
%如果uu=0，则p-i

%  
%

n = size(v,1); if v(i:n) == zeros(n - i + 1,1)  
n=尺寸（v，1）；如果v（i:n）=0（n-i+1,1）

P = eye(n); else  
P=眼睛（N）；其他

PP = eye(n - i + 1) - 2\*v(i:n)\*v(i:n)’;  
pp=眼睛（n-i+1）-2\*v（i:n）\*v（i:n）'；

P = [eye(i-1) zeros(i-1, n - i + 1); zeros(n - i + 1, i - 1) PP]; end end  
P=[眼（i-1）零（i-1，n-i+1）；零（n-i+1，i-1）p p]；结束

The function Hessenberg1 computes an upper Hessenberg matrix H and an orthogonal matrix Q such that A = Q>HQ.  
函数Hessenberg1计算一个上Hessenberg矩阵h和一个正交矩阵q，使a=q>hq。

function [H, Q] = Hessenberg1(A)  
函数[h，q]=Hessenberg1（a）

%  
%

% This function constructs an upper Hessenberg  
%此函数构造上Hessenberg

% matrix H and an orthogonal matrix Q such that  
%矩阵H和正交矩阵Q

% A = Q’ H Q  
%A=Q’H Q

n = size(A,1);  
n=尺寸（a，1）；

H = A;  
H＝a；

Q = eye(n); for i = 1:n-2  
Q=眼睛（n）；对于i=1:n-2

% H(i+1:n,i)  
%h（i+1:n，i）

[~,u] = house(H(i+1:n,i));  
[~，u]=房屋（h（i+1:n，i））；

% u  
%u

1. = buildhouse(u,1);  
   =建筑用房（U，1）；

Q(i+1:n,i:n) = P\*Q(i+1:n,i:n);  
q（i+1:n，i:n）=p\*q（i+1:n，i:n）；

H(i+1:n,i:n) = H(i+1:n,i:n) - 2\*u\*(u’)\*H(i+1:n,i:n);  
h（i+1:n，i:n）=h（i+1:n，i:n）-2\*u\*（u’）\*h（i+1:n，i:n）；

H(1:n,i+1:n) = H(1:n,i+1:n) - 2\*H(1:n,i+1:n)\*u\*(u’); end end  
H（1:N，I+1:N）=H（1:N，I+1:N）-2\*H（1:N，I+1:N）\*U\*（U’）；端部

Example 22.2. If  
例22.2。如果

,  
，

running Hessenberg1 we find  
运行Hessenberg1我们发现

|  |  |  |  |
| --- | --- | --- | --- |
|  1.0000 网络错误  H = −−50..38520000 网络错误   网络错误  0 网络错误  1.0000 网络错误 | −155..20693852 网络错误  −10..68930000 网络错误  − 网络错误  0 网络错误 | 0 网络错误  −10..68932069 网络错误  −0.0000 网络错误  0 网络错误 | 0  网络错误  −0.0000 网络错误  −0.0000 网络错误  0.0000 网络错误  0  网络错误 |

1. =  00 −00.8339.3714 −00.1516.5571 −−00..74285307.  
   =00−00.8339.3714−00.1516.5571−00..74285307。

  
γ

0 0.4082 −0.8165 0.4082  
0 0.4082−0.8165 0.4082

An important property of (upper) Hessenberg matrices is that if some subdiagonal entry Hp+1p = 0, then H is of the form  
Hessenberg矩阵（上）的一个重要性质是，如果某个次极性项hp+1p=0，则h的形式为

,  
，

where both H11 and H22 are upper Hessenberg matrices (with H11 a p×p matrix and H22 a (n − p) × (n − p) matrix), and the eigenvalues of H are the eigenvalues of H11 and H22. For  
其中h11和h22都是上Hessenberg矩阵（具有h11 a p×p矩阵和h22 a（n-p）×n-p矩阵），h的特征值是h11和h22的特征值。为了

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| example, in the matrix 网络错误 |  |  |  |  |
|  ∗ 网络错误  h21 网络错误  H =  0 网络错误   网络错误   0 网络错误  0 网络错误  if h43 = 0, then we have the block matrix 网络错误 | ∗ 网络错误  ∗ 网络错误  h32 网络错误  0 网络错误  0 网络错误 | ∗ 网络错误  ∗ 网络错误  h∗43 网络错误  0 网络错误 | ∗ 网络错误  ∗ 网络错误  ∗ 网络错误  ∗ h54 网络错误 | ∗ ∗ 网络错误  ∗, 网络错误  ∗ 网络错误  ∗ 网络错误 |
|  ∗ 网络错误  h21 网络错误  H =  0 网络错误   网络错误   0 网络错误  0 网络错误 | ∗ 网络错误  ∗ 网络错误  h32 网络错误  0 网络错误  0 网络错误 | ∗ 网络错误  ∗ 网络错误  ∗0 网络错误  0 网络错误 | ∗ 网络错误  ∗ ∗ 网络错误  ∗ h54 网络错误 | ∗ ∗ 网络错误  ∗. 网络错误  ∗ 网络错误  ∗ 网络错误 |

Then the list of eigenvalues of H is the concatenation of the list of eigenvalues of H11 and the list of the eigenvalues of H22. This is easily seen by induction on the dimension of the block H11.  
h的特征值列表是h11的特征值列表和h22的特征值列表的串联。通过对H11块尺寸的归纳，很容易看出这一点。

More generally, every upper Hessenberg matrix can be written in such a way that it has diagonal blocks that are Hessenberg blocks whose subdiagonal is not zero.  
更一般地说，每一个上海森堡矩阵都可以用这样的方式来写：它有对角块，这是海森堡块，其次对角块不是零。

Definition 22.2. An upper Hessenberg n × n matrix H is unreduced if hi+1i = 06 for i = 1,...,n − 1. A Hessenberg matrix which is not unreduced is said to be reduced.  
定义22.2.如果i=1，…，n−1的hi+1i=06，则上Hessenberg n×n矩阵h是未减少的。一个没有被约化的海森堡矩阵被称为约化矩阵。

The following is an example of an 8 × 8 matrix consisting of three diagonal unreduced Hessenberg blocks:  
以下是一个8×8矩阵的示例，该矩阵由三个对角未缩减的Hessenberg块组成：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | ? 网络错误  ? 网络错误  h32 网络错误  0 网络错误  0 网络错误  0 网络错误  0 网络错误  0 网络错误 | ? ? ? 网络错误  0 网络错误  0 网络错误  0 网络错误  0 网络错误  0 网络错误 | ∗ 网络错误  ∗ 网络错误  ∗? 网络错误  h54 网络错误  0 网络错误  0 网络错误  0 网络错误 | ∗ 网络错误  ∗ ∗? 网络错误  ? 网络错误  h65 网络错误  0 网络错误  0 网络错误 | ∗ 网络错误  ∗ ∗? 网络错误  ? 网络错误  ? 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  h87 网络错误 | ∗ 网络错误  ∗ 网络错误   ∗ 网络错误   网络错误  ∗. 网络错误   网络错误  ∗ 网络错误   网络错误  ∗ 网络错误  ? ? 网络错误 |

An interesting and important property of unreduced Hessenberg matrices is the following.  
下面是一个有趣的和重要的性质的未减少的海森堡矩阵。

Proposition 22.3. Let H be an n × n complex or real unreduced Hessenberg matrix. Then every eigenvalue of H is geometrically simple, that is, dim(Eλ) = 1 for every eigenvalue λ, where Eλ is the eigenspace associated with λ. Furthermore, if H is diagonalizable, then every eigenvalue is simple, that is, H has n distinct eigenvalues.  
提案22.3.设h为n×n复形或实无约Hessenberg矩阵。那么h的每个特征值在几何上都是简单的，即对于每个特征值λ，dim（eλ）=1，其中eλ是与λ相关的特征空间。此外，如果h是对角化的，那么每个特征值都是简单的，即h有n个不同的特征值。

Proof. We follow Serre’s proof [151] (Proposition 3.26). Let λ be any eigenvalue of H, let  
证据。我们遵循塞尔证明[151]（提案3.26）。设λ为h的任何特征值，设

M = λIn − H, and let N be the (1) matrix obtained from M by deleting its first row and its last column. Since is upper Hessenberg, N is a diagonal matrix with entries −hi+1i = 06 , i = 1,...,n − 1. Thus N is invertible and has rank n − 1. But a matrix has rank greater than or equal to the rank of any of its submatrices, so rank(M) = n − 1, since M is singular. By the rank-nullity theorem, rank(KerN) = 1, that is, dim(Eλ) = 1, as claimed.  
m=λin−h，n是从m中删除第一行和最后一列得到的（1）矩阵。由于是上赫森堡，n是一个对角线矩阵，条目−hi+1i=06，i=1，…，n−1。因此，n是可逆的，具有n-1的秩。但矩阵的秩大于或等于其任何子矩阵的秩，因此秩（m）=n-1，因为m是奇异的。根据秩零定理，秩（kern）=1，即dim（eλ）=1，如权利要求所述。

If H is diagonalizable, then the sum of the dimensions of the eigenspaces is equal to n, which implies that the eigenvalues of H are distinct.   
如果h是对角化的，那么特征空间的维数之和等于n，这意味着h的特征值是不同的。

As we said earlier, a case where Theorem 22.1 applies is the case where A is a symmetric (or Hermitian) positive definite matrix. This follows from two facts.  
如前所述，定理22.1适用的情况是，a是对称（或厄米特）正定矩阵的情况。这源于两个事实。

The first fact is that if A is Hermitian (or symmetric in the real case), then it is easy to show that the Hessenberg matrix similar to A is a Hermitian (or symmetric in real case) tridiagonal matrix. The conversion method is also more efficient. Here is an example of a symmetric tridiagonal matrix consisting of three unreduced blocks:  
第一个事实是，如果a是厄米特矩阵（或在实际情况下是对称的），那么很容易证明类似a的Hessenberg矩阵是厄米特矩阵（或在实际情况下是对称的）三对角矩阵。转换方法也更有效。下面是一个由三个未减少的块组成的对称三对角矩阵的示例：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | β1 网络错误  α2 β2 网络错误  0 网络错误  0 网络错误  0 网络错误  0 网络错误  0 网络错误 | 0 β2 网络错误  α3 网络错误  0 网络错误  0 网络错误  0 网络错误  0 网络错误  0 网络错误 | 0 网络错误  0 网络错误  0 α4 β4 网络错误  0 网络错误  0 网络错误  0 网络错误 | 0 网络错误  0 网络错误  0 β4 网络错误  α5 β5 网络错误  0 网络错误  0 网络错误 | 0 网络错误  0 网络错误  0 网络错误  0 β5 网络错误  α6 网络错误  0 网络错误  0 网络错误 | 0 网络错误  0 网络错误  0 网络错误  0 网络错误  0 网络错误  0 α7 β7 网络错误 | 0  网络错误  0  网络错误  0  网络错误  0  0 . 网络错误  0  网络错误  β7 α8 网络错误 |

Thus the problem of finding the eigenvalues of a symmetric (or Hermitian) matrix reduces to the problem of finding the eigenvalues of a symmetric (resp. Hermitian) tridiagonal matrix, and this can be done much more efficiently.  
因此，求对称（或厄米特）矩阵的特征值的问题可归结为求对称（或厄米特）矩阵的特征值的问题。赫密特）三对角矩阵，这可以更有效地完成。

The second fact is that if H is an upper Hessenberg matrix and if it is diagonalizable, then there is an invertible matrix P such that H = PΛP −1 with Λ a diagonal matrix consisting of the eigenvalues of H, such that P −1 has an LU-decomposition; see Serre [151] (Theorem  
第二个事实是，如果h是一个上海森堡矩阵，如果它是对角化的，那么就有一个可逆矩阵p，使得h=p∧p−1带有∧一个由h的特征值组成的对角矩阵，这样p−1有一个lu分解；见serre[151]（定理

13.3).  
13.3）。

As a consequence, since any symmetric (or Hermitian) tridiagonal matrix is a block diagonal matrix of unreduced symmetric (resp. Hermitian) tridiagonal matrices, by Proposition 22.3, we see that the QR algorithm applied to a tridiagonal matrix which is symmetric (or Hermitian) positive definite converges to a diagonal matrix consisting of its eigenvalues. Let us record this important fact.  
因此，由于任何对称（或厄米提亚）三对角矩阵都是非降阶对称（resp）的块对角矩阵。Hermitian）三对角矩阵，通过22.3，我们看到应用于对称（或Hermitian）正定的三对角矩阵的QR算法收敛到由其特征值组成的对角矩阵。让我们记录下这一重要事实。

Theorem 22.4. Let H be a symmetric (or Hermitian) positive definite tridiagonal matrix. If H is unreduced, then the QR algorithm converges to a diagonal matrix consisting of the eigenvalues of H.  
定理22.4.设h为对称（或厄米特）正定三对角矩阵。如果h不被约化，则qr算法收敛到由h的特征值组成的对角矩阵。

Since every symmetric (or Hermitian) positive definite matrix is similar to tridiagonal symmetric (resp. Hermitian) positive definite matrix, we deduce that we have a method for finding the eigenvalues of a symmetric (resp. Hermitian) positive definite matrix (more accurately, to find approximations as good as we want for these eigenvalues).  
因为每一个对称（或厄米特）正定矩阵都与三对角对称（resp）相似。厄米特）正定矩阵，我们推导出一种求对称（resp）特征值的方法。正定矩阵（更准确地说，为了找到这些特征值的近似值）。

If A is a symmetric (or Hermitian) matrix, since its eigenvalues are real, for some µ > 0 large enough (pick µ > ρ(A)), A + µI is symmetric (resp. Hermitan) positive definite, so we can apply the QR algorithm to an upper Hessenberg matrix similar to A+µI to find its eigenvalues, and then the eigenvalues of A are obtained by subtracting µ.  
如果a是对称（或厄米特）矩阵，由于其特征值是实的，对于一些足够大的μ>0（pickμ>ρ（a）），a+μi是对称的（resp.Hermitan）正定的，因此我们可以将qr算法应用到一个类似于a+μi的上Hessenberg矩阵中，找到它的特征值，然后通过减去μ得到a的特征值。

The problem of finding the eigenvalues of a symmetric matrix is discussed extensively in Parlett [131], one of the best references on this topic.  
关于对称矩阵的特征值的求法问题，帕莱特[131]对此作了广泛的讨论，这是本课题的最佳参考文献之一。

The upper Hessenberg form also yields a way to handle singular matrices. First, checking the proof of Proposition 13.21 that an n × n complex matrix A (possibly singular) can be factored as A = QR where Q is a unitary matrix which is a product of Householder reflections and R is upper triangular, it is easy to see that if A is upper Hessenberg, then Q is also upper Hessenberg. If H is an unreduced upper Hessenberg matrix, since Q is upper Hessenberg and R is upper triangular, we have hi+1i = qi+1irii for i = 1...,n−1, and since H is unreduced, rii = 06 for i = 1,...,n−1. Consequently H is singular iff rnn = 0. Then the matrix RQ is a matrix whose last row consists of zero’s thus we can deflate the problem by considering the (n − 1) × (n − 1) unreduced Hessenberg matrix obtained by deleting the last row and the last column. After finitely many steps (not larger that the multiplicity of the eigenvalue 0), there remains an invertible unreduced Hessenberg matrix. As an alternative, see Serre [151] (Chapter 13, Section 13.3.2).  
上海森堡形式也产生了一种处理奇异矩阵的方法。首先，检查命题13.21的证明，一个n×n的复数矩阵a（可能是奇异的）可以被分解为a=q r，其中q是户主反射的乘积，r是上三角形，很容易看出，如果a是上Hessenberg，那么q也是上Hes森伯格。如果h是一个未减少的上Hessenberg矩阵，因为q是上Hessenberg矩阵，r是上三角形，我们有h i+1i=qi+1i rii表示i=1…，n-1，并且由于h是未减少的，rii=06表示i=1，…，n-1。因此，h是单数iff rnn=0。那么矩阵rq是最后一行由零组成的矩阵，因此我们可以通过考虑删除最后一行和最后一列得到的（n-1）×（n-1）未减少的Hessenberg矩阵来消除问题。在有限多个步骤之后（不大于特征值0的多重性），仍然存在一个可逆的不可约Hessenberg矩阵。作为替代方案，见SERRE[151]（第13章，第13.3.2节）。

As is, the QR algorithm, although very simple, is quite inefficient for several reasons. In the next section, we indicate how to make the method more efficient. This involves a lot of work and we only discuss the main ideas at a high level.  
事实上，QR算法虽然很简单，但由于几个原因效率很低。在下一节中，我们将说明如何提高方法的效率。这涉及到很多工作，我们只在高层讨论主要想法。

## 22.3 Making the QR Method More Efficient Using Shifts 22.3提高使用轮班的QR方法的效率

To improve efficiency and cope with pairs of complex conjugate eigenvalues in the case of real matrices, the following steps are taken:  
为了提高效率，并应对实矩阵中的复共轭特征值对，采取以下步骤：

1. Initially reduce the matrix A to upper Hessenberg form, as A = UHU∗. Then apply the QR-algorithm to H (actually, to its unreduced Hessenberg blocks). It is easy to see that the matrices Hk produced by the QR algorithm remain upper Hessenberg.  
   最初将矩阵A简化为上赫森堡形式，即A=uhu。然后将qr算法应用于h（实际上，应用于其未减少的hessenberg块）。很容易看出，由QR算法生成的矩阵hk仍保持在Hessenberg的上方。
2. To accelerate convergence, use shifts, and to deal with pairs of complex conjugate eigenvalues, use double shifts.  
   为了加速收敛，使用移位，为了处理复共轭特征值对，使用双移位。
3. Instead of computing a QR-factorization explicitly while doing a shift, perform an implicit shift which computes without having to compute a QRfactorization (of Ak − σkI), and similarly in the case of a double shift. This is the most intricate modification of the basic QR algorithm and we will not discuss it here. This method is usually referred as bulge chasing. Details about this technique for real matrices can be found in Demmel [49] (Section 4.4.8) and Golub and Van Loan [80] (Section 7.5). Watkins discusses the QR algorithm with shifts as a bulge chasing method in the more general case of complex matrices [181, 182].  
   在进行移位时，不需要显式计算qr因子分解，而是执行隐式移位，该移位不需要计算qr factorization（ak-σki），同样，在进行双移位时也需要计算qrfactorization。这是对基本QR算法最复杂的修改，我们在这里不讨论它。这种方法通常被称为鼓包追逐。有关真实矩阵的这种技术的详细信息，请参见demmel[49]（第4.4.8节）和Golub和van Loan[80]（第7.5节）。在复杂矩阵的更一般情况下，Watkins讨论了将移位作为凸点追踪方法的QR算法[181182]。

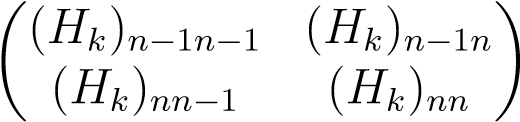
Let us repeat an important remark made in the previous section. If we start with a matrix H in upper Hessenberg form, if at any stage of the QR algorithm we find that some subdiagonal entry (Hk)p+1p = 0 or is very small, then Hk is of the form  
让我们重复上一节中的一个重要评论。如果我们以Hessenberg上形式的矩阵h开始，如果在qr算法的任何阶段，我们发现某些次方向项（hk）p+1p=0或非常小，那么hk就是这种形式。

,  
，

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| a (n − p) × (n − p) matrix), and the eigenvalues of Hk are the eigenvalues of H11 and H22. 网络错误 | | | | |
| For example, in the matrix 网络错误 |  |  |  |  |
|  ∗ 网络错误  h21 网络错误  H =  0 网络错误   网络错误   0 网络错误  0 网络错误  if h43 = 0, then we have the block matrix 网络错误 | ∗ 网络错误  ∗ 网络错误  h32 网络错误  0 网络错误  0 网络错误 | ∗ 网络错误  ∗ 网络错误  h∗43 网络错误  0 网络错误 | ∗ 网络错误  ∗ 网络错误  ∗ 网络错误  ∗ h54 网络错误 | ∗ ∗ 网络错误  ∗, 网络错误  ∗ 网络错误  ∗ 网络错误 |
|  ∗ 网络错误  h21 网络错误  H =  0 网络错误   网络错误   0 网络错误  0 网络错误 | ∗ 网络错误  ∗ 网络错误  h32 网络错误  0 网络错误  0 网络错误 | ∗ 网络错误  ∗ 网络错误  ∗0 网络错误  0 网络错误 | ∗ 网络错误  ∗ ∗ 网络错误  ∗ h54 网络错误 | ∗ ∗ 网络错误  ∗. 网络错误  ∗ 网络错误  ∗ 网络错误 |

where both H11 and H22 are upper Hessenberg matrices (with H11 a p × p matrix and H22 Then we can recursively apply the QR algorithm to H11 and H22.  
其中h11和h22都是上Hessenberg矩阵（h11是p×p矩阵和h22），那么我们可以递归地将qr算法应用于h11和h22。

In particular, if (Hk)nn−1 = 0 or is very small, then (Hk)nn is a good approximation of an eigenvalue, so we can delete the last row and the last column of Hk and apply the QR algorithm to this submatrix. This process is called deflation. If (Hk)n−1n−2 = 0 or is very small, then the 2 × 2 “corner block”  
特别是，如果（hk）nn−1=0或非常小，那么（hk）nn是特征值的良好近似值，因此我们可以删除hk的最后一行和最后一列，并将qr算法应用于该子矩阵。这个过程叫做通货紧缩。如果（hk）n−1n−2=0或非常小，则2×2“角块”



appears, and its eigenvalues can be computed immediately by solving a quadratic equation. Then we deflate Hk by deleting its last two rows and its last two columns and apply the QR algorithm to this submatrix.  
出现了，它的特征值可以通过解二次方程立即计算出来。然后，我们通过删除最后两行和最后两列来缩小hk，并将qr算法应用于该子矩阵。

Thus it would seem desirable to modify the basic QR algorithm so that the above situations arises, and this is what shifts are designed for. More precisely, under the hypotheses of Theorem 22.1, it can be shown (see Ciarlet [41], Section 6.3) that the entry (Ak)ij with i > j converges to 0 as |λi/λj|k converges to 0. Also, if we let ri be defined by  
因此，似乎需要修改基本的QR算法，以便出现上述情况，这就是移位的设计目的。更准确地说，在定理22.1的假设下，可以证明（见Ciarlet[41]第6.3节），i>j的入口（ak）ij收敛到0，因为λi/λj\_k收敛到0。另外，如果我们让ri定义为

,  
，

then there is a constant C (independent of k) such that  
然后有一个常数c（独立于k），这样



In particular, if H is upper Hessenberg, then the entry (Hk)i+1i converges to 0 as |λi+1/λi|k converges to 0. Thus if we pick σk close to λi, we expect that (Hk − σkI)i+1i converges to 0 as |(λi+1 −σk)/(λi −σk)|k converges to 0, and this ratio is much smaller than 1 as σk is closer to λi. Typically, we apply a shift to accelerate convergence to λn (so i = n − 1). In this case, both (Hk −σkI)nn−1 and |(Hk −σkI)nn −λn| converge to 0 as |(λn −σk)/(λn−1 −σk)|k converges to 0.  
特别是，如果h是上Hessenberg，则条目（h k）i+1i收敛到0，因为λi+1/λi\_k收敛到0。因此，如果我们选取靠近λi的σk，我们期望（hk−σki）i+1i收敛到0，因为（λi+1−σk）/（λi−σk）k收敛到0，并且这个比率比1小得多，因为σk更接近λi。通常，我们应用移位来加速收敛到λn（因此i=n-1）。在这种情况下，（hk−σki）nn-1和（hk−σki）nn−λn收敛到0，因为（λn−σk）/（λn−1−σk）k收敛到0。

A shift is the following modified QR-steps (switching back to an arbitrary matrix A, since the shift technique applies in general). Pick some σk, hopefully close to some eigenvalue of A (in general, λn), and QR-factor Ak − σkI as  
移位是以下修改的QR步骤（切换回任意矩阵A，因为移位技术通常适用）。选择一些σk，希望接近a的特征值（一般来说，λn），以及qr因子ak-σki作为

Ak − σkI = QkRk,  
AK−σki=qkrk，

and then form  
然后形成

Ak+1 = RkQk + σkI.  
AK+1=RKQK+σki。

Since  
自从

Ak+1 = RkQk + σkI  
AK+1=RKQK+σki

= Q∗kQkRkQk + Q∗kQkσk  
=q kqkrkqk+q k q kσk

= Q∗k(QkRk + σkI)Qk  
=q k（qkrk+σki）qk

= Q∗kAkQk,  
=Q Kakqk，

Ak+1 is similar to Ak, as before. If Ak is upper Hessenberg, then it is easy to see that Ak+1 is also upper Hessenberg.  
AK+1与AK类似，如前所述。如果ak是上海森堡，那么很容易看出ak+1也是上海森堡。

If A is upper Hessenberg and if σi is exactly equal to an eigenvalue, then Ak − σkI is singular, and forming the QR-factorization will detect that Rk has some diagonal entry equal to 0. Assuming that the QR-algorithm returns (Rk)nn = 0 (if not, the argument is easily adapted), then the last row of RkQk is 0, so the last row of Ak+1 = RkQk + σkI ends with σk (all other entries being zero), so we are in the case where we can deflate Ak (and σk is indeed an eigenvalue).  
如果a是上Hessenberg，如果σi正好等于特征值，那么ak-σki是奇异的，形成qr因式分解将检测到Rk有一些等于0的对角线入口。假设qr算法返回（rk）nn=0（如果不是，参数很容易适应），那么RKQK的最后一行是0，所以AK+1=RKQK+σki的最后一行以σk结尾（所有其他项都为零），所以我们可以对ak进行放气（而σk实际上是一个特征值）。

The question remains, what is a good choice for the shift σk?  
问题是，对于移位σk，什么是一个好的选择？

Assuming again that H is in upper Hessenberg form, it turns out that when (Hk)nn−1 is small enough, then a good choice for σk is (Hk)nn. In fact, the rate of convergence is quadratic, which means roughly that the number of correct digits doubles at every iteration. The reason is that shifts are related to another method known as inverse iteration, and such a method converges very fast. For further explanations about this connection, see Demmel [49] (Section 4.4.4) and Trefethen and Bau [171] (Lecture 29).  
再次假设h为上Hessenberg形式，结果表明当（h k）nn-1足够小时，σk的一个好选择是（hk）nn。事实上，收敛速度是二次的，这意味着在每次迭代中，正确数字的数目大致都会加倍。原因是移位与另一种称为逆迭代的方法有关，这种方法收敛得很快。关于这种联系的进一步解释，见demmel[49]（第4.4.4节）和trefethen和bau[171]（第29课）。

One should still be cautious that the QR method with shifts does not necessarily converge, and that our convergence proof no longer applies, because instead of having the identity Ak = PkRk, we have  
我们仍然应该谨慎，带移位的qr方法不一定收敛，并且我们的收敛证明不再适用，因为我们没有拥有标识ak=pkrk，而是

(A − σkI)···(A − σ2I)(A − σ1I) = PkRk.  
（a−σki）···（a−σ2i）（a−σ1i）=pkrk。

Of course, the QR algorithm loops immediately when applied to an orthogonal matrix A. This is also the case when A is symmetric but not positive definite. For example, both the QR algorithm and the QR algorithm with shifts loop on the matrix  
当然，当应用于正交矩阵A时，QR算法会立即循环，当A是对称的但不是正定的时候也是如此。例如，QR算法和矩阵上带移位环的QR算法

.  
.

In the case of symmetric matrices, Wilkinson invented a shift which helps the QR algorithm with shifts to make progress. Again, looking at the lower corner of Ak, say  
在对称矩阵的情况下，威尔金森发明了一种移位，这有助于QR算法的移位取得进展。再看看AK的下角，说

,  
，

the Wilkinson shift picks the eigenvalue of B closer to an. If we let  
威尔金森位移选取的特征值B更接近A。如果我们让

δ = an−1 − an ,  
δ=an−1−an，

2  
二

it is easy to see that the eigenvalues of B are given by  
很容易看出，b的特征值由

.  
.

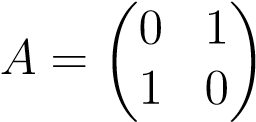
It follows that  
接下来是

q λ − an = δ ± δ2 + b2n−1,  
qλ−an=δ±δ2+b2n−1，

and from this it is easy to see that the eigenvalue closer to an is given by  
从这个很容易看出，更接近a的特征值是由

.  
.

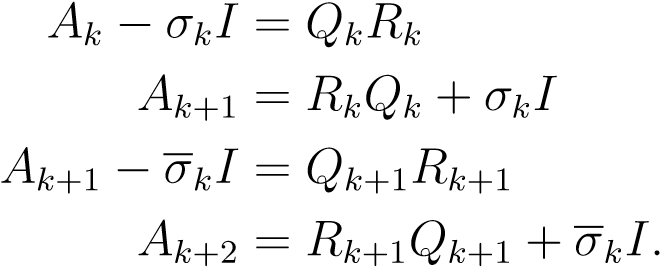
If δ = 0, then we pick arbitrarily one of the two eigenvalues. Observe that the Wilkinson shift applied to the matrix  
如果δ=0，那么我们可以任意选取两个特征值中的一个。观察应用于矩阵的威尔金森位移



is either +1 or −1, and in one step, deflation occurs and the algorithm terminates successfully.  
是+1或−1，在一个步骤中，会发生通缩，算法成功终止。

We now discuss double shifts, which are intended to deal with pairs of complex conjugate eigenvalues.  
我们现在讨论双移位，这是为了处理复共轭特征值对。

Let us assume that A is a real matrix. For any complex number σk with nonzero imaginary part, a double shift consists of the following steps:  
假设A是一个实矩阵。对于具有非零虚数部分的复数σk，双移位包括以下步骤：



From the computation made for a single shift, we have Ak+1 = Q∗kAkQk and Ak+2 =  
根据对单个位移的计算，我们得到了ak+1=q kakqk和ak+2。=

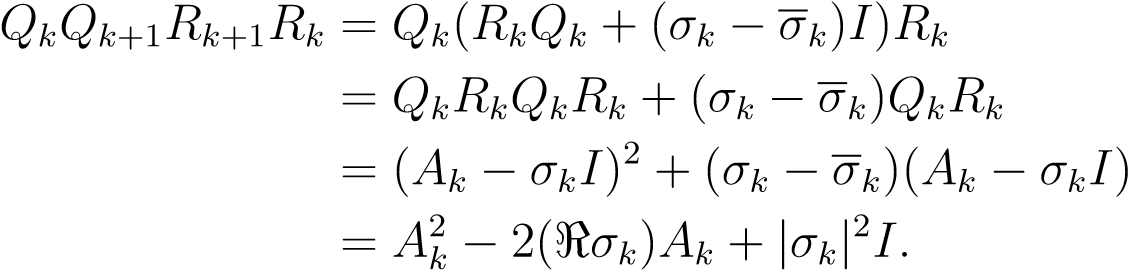
Q∗k+1Ak+1Qk+1, so we obtain  
q k+1ak+1qk+1，因此我们得出

.  
.

The matrices Qk are complex, so we would expect that the Ak are also complex, but remarkably we can keep the products QkQk+1 real, and so the Ak also real. This is highly desirable to avoid complex arithmetic, which is more expensive. Observe that since  
矩阵qk是复杂的，所以我们希望ak也是复杂的，但很明显我们可以保持产品qkqk+1是真实的，所以ak也是真实的。这是非常理想的避免复杂的算术，这是更昂贵的。从那以后再观察

Qk+1Rk+1 = Ak+1 − σkI = RkQk + (σk − σk)I,  
qk+1rk+1=ak+1−σk i=rkqk+（σk−σk）i，

we have  
我们有



If we assume by induction that matrix Ak is real (with k = 2`+1,` ≥ 0), then the matrix S = A2k − 2(<σk)Ak + |σk|2I is also real, and since QkQk+1 is unitary and Rk+1Rk is upper triangular, we see that  
如果我们通过归纳假设矩阵ak是实的（k=2`+1，`≥0），那么矩阵s=a2k−2（<σk）ak+σk 2i也是实的，因为qkqk+1是一元的，而rk+1rk是上三角的，我们可以看到

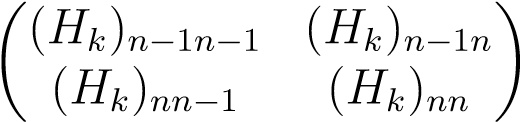
S = QkQk+1Rk+1Rk  
S=QKQK+1RK+1RK

is a QR-factorization of the real matrix S, thus QkQk+1 and Rk+1Rk can be chosen to be real matrices, in which case (QkQk+1)∗ is also real, and thus  
是实矩阵s的qr因子分解，因此qkqk+1和rk+1rk可以选择为实矩阵，在这种情况下（qkqk+1）也是实矩阵，因此

Ak+2 = Q∗k+1Q∗kAkQkQk+1 = (QkQk+1)∗AkQkQk+1  
ak+2=q k+1q kakqk+1=（qkqk+1）akqkqk+1

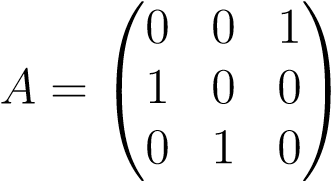
is real. Consequently, if A1 = A is real, then A2`+1 is real for all ` ≥ 0.  
是真的。因此，如果a1=a是实的，那么a2`+1对所有的`≥0都是实的。

The strategy that consists in picking σk and σk as the complex conjugate eigenvalues of the corner block  
选择σk和σk作为角块复共轭特征值的策略



is called the Francis shift (here we are assuming that A has be reduced to upper Hessenberg form).  
被称为弗朗西斯位移（这里我们假设a已经被简化为上海森堡形式）。

It should be noted that there are matrices for which neither a shift by (Hk)nn nor the Francis shift works. For instance, the permutation matrix  
应该注意的是，有些矩阵的（hk）nn移位和Francis移位都不起作用。例如，置换矩阵



has eigenvalues ei2π/3,ei4π/3,+1, and neither of the above shifts apply to the matrix  
具有特征值ei2π/3，ei4π/3，+1，并且上述两种移位都不适用于矩阵

.  
.

However, a shift by 1 does work. There are other kinds of matrices for which the QR algorithm does not converge. Demmel gives the example of matrices of the form  
但是，按1的移位确实有效。还有其他类型的矩阵，QR算法不收敛。demmel给出了形式矩阵的例子。

0 1 0 0  
0 1 0 0\_

1 0 h 0  
10小时0\_

0 −h 0 1  
0−H 0 1

0 0 1 0  
0 0 1 0

where h is small.  
其中h是小的。

Algorithms implementing the QR algorithm with shifts and double shifts perform “exceptional” shifts every 10 shifts. Despite the fact that the QR algorithm has been perfected since the 1960’s, it is still an open problem to find a shift strategy that ensures convergence of all matrices.  
采用移位和双移位的QR算法每10个移位执行一次“异常”移位。尽管自20世纪60年代以来QR算法得到了完善，但寻找一种确保所有矩阵收敛的移位策略仍然是一个开放性问题。

Implicit shifting is based on a result known as the implicit Q theorem. This theorem says that if A is reduced to upper Hessenberg form as A = UHU∗ and if H is unreduced (hi+1i = 06 for i = 1,...,n−1), then the columns of index 2,...,n of U are determined by the first column of U up to sign; see Demmel [49] (Theorem 4.9) and Golub and Van Loan [80] (Theorem 7.4.2) for the proof in the case of real matrices. Actually, the proof is not difficult and will be the object of a homework exercise. In the case of a single shift, an implicit shift generates Ak+1 = Q∗kAkQk without having to compute a QR-factorization of Ak − σkI. For real matrices, this is done by applying a sequence of Givens rotations which perform a bulge chasing process (a Givens rotation is an orthogonal block diagonal matrix consisting of a single block which is a 2D rotation, the other diagonal entries being equal to 1). Similarly, in the case of a double shift, Ak+2 = (QkQk+1)∗AkQkQk+1 is generated without having to compute the QR-factorizations of Ak − σkI and Ak+1 − σkI. Again, (QkQk+1)∗AkQkQk+1 is generated by applying some simple orthogonal matrices which perform a bulge chasing process. See Demmel [49] (Section 4.4.8) and Golub and Van Loan [80] (Section 7.5) for further explanations regarding implicit shifting involving bulge chasing in the case of real matrices. Watkins [181, 182] discusses bulge chasing in the more general case of complex matrices.  
隐式移位是基于一个被称为隐式Q定理的结果。这个定理表明，如果a被简化为a=u h u的上海森堡形式，而h未被简化（i=1，…，n−1，hi+1i=06），那么索引2，…，n的u列由u的第一列直到符号决定；见demmel[49]（定理4.9）和golub和van loan[80]（定理7.4.2）。对于实矩阵的证明。事实上，证明并不难，而且将是家庭作业练习的对象。在单个移位的情况下，隐式移位生成ak+1=q kakqk，而无需计算ak-σki的qr因子分解。对于实矩阵，这是通过应用一系列执行凸起追踪过程的givens旋转来完成的（givens旋转是一个正交的块对角矩阵，由一个二维旋转的单个块组成，其他对角线条目等于1）。同样，在双移位的情况下，AK+2=（QKQK+1）AKQKQK+1生成时不需要计算AK−σki和AK+1−σki的QR因子分解。同样地，（qkqk+1）akqkqk+1是通过应用一些简单的正交矩阵来生成的，这些矩阵执行一个凸起追踪过程。参见demmel[49]（第4.4.8节）和Golub和van Loan[80]（第7.5节），了解关于真实矩阵中涉及凸起追踪的隐式移位的进一步解释。Watkins[181182]讨论了复杂矩阵的更一般情况下的凸度追踪。

The Matlab function for finding the eigenvalues and the eigenvectors of a matrix A is eig and is called as [U, D] = eig(A). It is implemented using an optimized version of the QR-algorithm with implicit shifts.  
求矩阵A的特征值和特征向量的matlab函数是特征值，称为[u，d]=eig（a）。它是使用隐式移位的优化版QR算法实现的。

If the dimension of the matrix A is very large, we can find approximations of some of the eigenvalues of A by using a truncated version of the reduction to Hessenberg form due to Arnoldi in general and to Lanczos in the symmetric (or Hermitian) tridiagonal case.  
如果矩阵A的维数非常大，我们可以通过使用截断形式的约简来找到A的一些特征值的近似值，这种约简形式通常是由于阿诺迪和兰佐斯在对称（或厄米提亚）三对角情况下的约简。

## 22.4 Krylov Subspaces; Arnoldi Iteration 22.4 Krylov子空间；Arnoldi迭代

In this section, we denote the dimension of the square real or complex matrix A by m rather than n, to make it easier for the reader to follow Trefethen and Bau exposition [171], which is particularly lucid.  
在这一节中，我们用m而不是n来表示正方形实矩阵或复矩阵的维数，以便读者更容易遵循Trefetten和Bau论述[171]，这一点特别清晰。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| upper left block 网络错误 |  |  |  |  |  |
|  |  网络错误  h11 网络错误  h21 网络错误   网络错误   0 网络错误  Hen =  ... 网络错误    0 网络错误   网络错误  0 网络错误 | h12 h22 h32 ... 网络错误  ··· 网络错误  ··· 网络错误 | h13 h23 h33 ... 网络错误  0 网络错误  0 网络错误 | ··· ··· 网络错误  ···... 网络错误  hnn−1 网络错误  0 网络错误 | h1n  h2n  h3n  网络错误  ...  网络错误  hnn  hn+1n 网络错误 |

Suppose that the m × m matrix A has been reduced to the upper Hessenberg form H, as A = UHU∗. For any n ≤ m (typically much smaller than m), consider the (n + 1) × n  
假设m×m矩阵a已简化为上赫森堡形式h，即a=uhu。对于任何n≤m（通常小于m），考虑（n+1）×n

22.4. KRYLOV SUBSPACES; ARNOLDI ITERATION of H, and the n × n upper Hessenberg matrix Hn obtained by deleting the last row of Hen,  
22.4。Krylov子空间；H的Arnoldi迭代，以及通过删除最后一行hen得到的n×n上Hessenberg矩阵hn，

.  
.

If we denote by Un the m×n matrix consisting of the first n columns of U, denoted u1,...,un, then matrix consisting of the first n columns of the matrix UH = AU can be expressed as  
如果用u n表示由u的前n列组成的m×n矩阵，表示为u1，…，un，那么由矩阵的前n列组成的矩阵uh=au可以表示为

AUn = Un+1Hen. (∗1)  
aun=un+1小时。（1）

It follows that the nth column of this matrix can be expressed as  
因此，该矩阵的第n列可以表示为

Aun = h1nu1 + ··· + hnnun + hn+1nun+1. (∗2)  
aun=h1nu1+····+hnnun+hn+1nun+1。（2）

Since (u1,...,un) form an orthonormal basis, we deduce from (∗2) that  
由于（u1，…，un）形成了正态基，我们从（2）推导出

huj,Auni = u∗jAun = hjn, j = 1,...,n. (∗3)  
huj，auni=u\_jaun=hjn，j=1，…，n.（3）

Equations (∗2) and (∗3) show that Un+1 and Hen can be computed iteratively using the following algorithm due to Arnoldi, known as Arnoldi iteration:  
方程（2）和（3）表明，由于Arnoldi（称为Arnoldi迭代），可以使用以下算法迭代计算un+1和hen：

Given an arbitrary nonzero vector b ∈ Cm, let u1 = b/kbk; for n = 1,2,3,... do z := Aun; for j = 1 to n do hjn := u∗jz;  
给定任意非零向量b∈cm，设u1=b/kbk；对于n=1,2,3，…do z：=aun；对于j=1至n do hjn：=u jz；

z := z − hjnuj  
Z：=Z−HJnuj

endfor hn+1n := kzk; if hn+1n = 0 quit un+1 = z/hn+1n  
endfor hn+1n：=kzk；如果hn+1n=0退出un+1=z/hn+1n

When hn+1n = 0, we say that we have a breakdown of the Arnoldi iteration.  
当hn+1n=0时，我们说Arnoldi迭代有一个分解。

Arnoldi iteration is an algorithm for producing the n×n Hessenberg submatrix Hn of the full Hessenberg matrix H consisting of its first n rows and n columns (the first n columns of U are also produced), not using Householder matrices.  
Arnoldi迭代是一种生成完整的Hessenberg矩阵h的n×n Hessenberg子矩阵hn的算法，该矩阵由其前n行和n列组成（U的前n列也是生成的），而不使用户主矩阵。

As long as hj+1j = 06 for j = 1,...,n, Equation (∗2) shows by an easy induction that un+1 belong to the span of (b,Ab,...,Anb), and obviously Aun belongs to the span of (u1,...,un+1), and thus the following spaces are identical:  
只要j=1，…，n的hj+1j=06，方程（2）通过一个简单的归纳表明，un+1属于（b，ab，…，anb）的跨度，而aun显然属于（u1，…，un+1）的跨度，因此以下空间是相同的：

Span(b,Ab,...,Anb) = Span(u1,...,un+1).  
span（b，ab，…，anb）=span（u1，…，un+1）。

The space Kn(A,b) = Span(b,Ab,...,An−1b) is called a Krylov subspace. We can view Arnoldi’s algorithm as the construction of an orthonormal basis for Kn(A,b). It is a sort of Gram–Schmidt procedure.  
空间kn（a，b）=span（b，ab，…，an−1b）称为krylov子空间。我们可以将Arnoldi算法看作是构造kn（a，b）的正交基。这是一种克-施密特程序。

Equation (∗2) shows that if Kn is the m × n matrix whose columns are the vectors (b,Ab,...,An−1b), then there is a n × n upper triangular matrix Rn such that  
方程（2）表明，如果kn是m×n矩阵，其列为向量（b，ab，…，an−1b），则存在n×n上三角矩阵rn，从而

Kn = UnRn. (∗4)  
kn=unrn.（4）

The above is called a reduced QR factorization of Kn.  
上面称为kn的简化qr因子分解。

Since (u1,...,un) is an orthonormal system, the matrix is the +1) matrix consisting of the identity matrix In plus an extra column of 0’s, so is  
因为（u1，…，un）是一个正交系统，矩阵是由单位矩阵加上一个0的额外列组成的+1）矩阵，所以

obtained by deleting the last row of Hen, namely Hn, and so  
通过删除母鸡的最后一行，即hn获得，依此类推。

. (∗5)  
.（5）

We summarize the above facts in the following proposition.  
我们将上述事实概括为以下命题。

Proposition 22.5. If Arnoldi iteration run on an m × m matrix A starting with a nonzero vector b ∈ Cm does not have a breakdown at stage n ≤ m, then the following properties hold:  
提案22.5.如果Arnoldi迭代在m×m矩阵a上以非零向量b∈cm开始运行，在n≤m阶段没有崩溃，那么以下属性保持不变：

1. If Kn is the m × n Krylov matrix associated with the vectors (b,Ab,...,An−1b) and if Un is the m × n matrix of orthogonal vectors produced by Arnoldi iteration, then there is a QR-factorization  
   如果kn是与向量（b，ab，…，a n−1b）相关联的m×n krylov矩阵，如果un是由Arnoldi迭代生成的正交向量的m×n矩阵，则存在qr因子分解。

Kn = UnRn,  
kn=unrn，

for some n × n upper triangular matrix Rn.  
对于一些n×n上三角矩阵rn。

1. The m×n upper Hessenberg matrices Hn produced by Arnoldi iteration are the projection of A onto the Krylov space Kn(A,b), that is,  
   Arnoldi迭代产生的m×n上Hessenberg矩阵hn是a对krylov空间kn（a，b）的投影，也就是说，

.  
.

1. The successive iterates are related by the formula  
   连续迭代与公式有关

.  
.

Remark: If Arnoldi iteration has a breakdown at stage n, that is, hn+1 = 0, then we found the first unreduced block of the Hessenberg matrix H. It can be shown that the eigenvalues of Hn are eigenvalues of A. So a breakdown is actually a good thing. In this case, we can pick some new nonzero vector un+1 orthogonal to the vectors (u1,...,un) as a new starting vector and run Arnoldi iteration again. Such a vector exists since the (n+1)th column of U works. So repeated application of Arnoldi yields a full Hessenberg reduction of A. However,  
注：如果Arnoldi迭代在n阶段有一个分解，即hn+1=0，那么我们就找到了Hessenberg矩阵h的第一个未简化块，可以证明hn的特征值是a的特征值，所以分解实际上是一件好事。在这种情况下，我们可以选择一些与向量（u1，…，un）正交的新的非零向量un+1作为新的起始向量，并再次运行arnoldi迭代。这种向量存在于u的第（n+1）列工作之后。因此，重复使用阿诺迪得到了一个完整的海森堡减少a。

### 22.4. KRYLOV SUBSPACES; ARNOLDI ITERATION 22.4。Krylov子空间；Arnoldi迭代

this is not what we are after, since m is very large an we are only interested in a “small” number of eigenvalues of A.  
这不是我们所追求的，因为m非常大，我们只对a的“少量”特征值感兴趣。

There is another aspect of Arnoldi iteration, which is that it solves an optimization problem involving polynomials of degree n. Let Pn denote the set of (complex) monic polynomials of degree n, that is, polynomials of the form  
Arnoldi迭代还有一个方面，它解决了一个涉及n次多项式的优化问题。让pn表示n次（复数）多项式的集合，即形式的多项式。

p(z) = zn + cn−1zn−1 + ··· + c1z + c0 (ci ∈ C).  
p（z）=zn+cn−1zn−1+····+c1z+c0（ci∈c）。

For any m × m matrix A, we write  
对于任何M×M矩阵A，我们写

p(A) = An + cn−1An−1 + ··· + c1A + c0I.  
p（a）=an+cn−1an−1+····+c1a+c0i。

The following result is proven in Trefethen and Bau [171] (Lecture 34, Theorem 34.1).  
以下结果在Trefethen和Bau[171]中得到了证明（第34课，定理34.1）。

Theorem 22.6. If Arnoldi iteration run on an m × m matrix A starting with a nonzero vector b does not have a breakdown at stage n ≤ m, then there is a unique polynomial p ∈ Pn such that kp(A)bk2 is minimum, namely the characteristic polynomial det(zI − Hn) of Hn.  
定理22.6。如果以非零向量b开始的m×m矩阵a上运行的Arnoldi迭代在n≤m阶段没有崩溃，则存在一个唯一的多项式p∈pn，使得kp（a）bk2最小，即hn的特征多项式det（zi-hn）。

Theorem 22.6 can be viewed as the “justification” for a method to find some of the eigenvalues of of them). Intuitively, the closer the roots of the characteristic polynomials of Hn are to the eigenvalues of A, the smaller kp(A)bk2 should be, and conversely. In the extreme case where m = n, by the Cayley–Hamilton theorem, p(A) = 0 (where p is the characteristic polynomial of A), so this idea is plausible, but this is far from constituting a proof (also, b should have nonzero coordinates in all directions associated with the eigenvalues).  
定理22.6可被视为一种求其某些特征值的方法的“正当性”。直观地说，hn特征多项式的根越接近a的特征值，kp（a）bk2越小，反之亦然。在m=n的极端情况下，根据凯莱-汉密尔顿定理，p（a）=0（其中p是a的特征多项式），所以这个想法是合理的，但这远不能构成一个证明（同时，b在与特征值相关的所有方向上都应该有非零坐标）。

The method known as the Rayleigh–Ritz method is to run Arnoldi iteration on A and some b = 06 chosen at random for steps before or until a breakdown occurs. Then run the QR algorithm with shifts on Hn. The eigenvalues of the Hessenberg matrix Hn may then be considered as approximations of the eigenvalues of A. The eigenvalues of Hn are called Arnoldi estimates or Ritz values. One has to be cautious because Hn is a truncated version of the full Hessenberg matrix H, so not all of the Ritz values are necessary close to eigenvalues of A. It has been observed that the eigenvalues that are found first are the extreme eigenvalues of A, namely those close to the boundary of the spectrum of A plotted in C. So if A has real eigenvalues, the largest and the smallest eigenvalues appear first as Ritz values. In many problems where eigenvalues occur, the extreme eigenvalues are the one that need to be computed. Similarly, the eigenvectors of Hn may be considered as approximations of eigenvectors of A.  
称为瑞利-里兹方法的方法是在a和一些b=06上运行阿诺迪迭代，随机选择步骤，直到出现故障。然后在hn上运行移位的qr算法。然后，可以将Hessenberg矩阵hn的特征值视为a特征值的近似值。hn的特征值称为Arnoldi估计或Ritz值。我们必须谨慎，因为hn是完整的Hessenberg矩阵h的截尾形式，所以并非所有的Ritz值都必须接近a的特征值。据观察，首先找到的特征值是a的极端特征值，即那些接近a的边界的特征值。图中A的谱，如果A有实特征值，最大和最小的特征值首先作为Ritz值出现。在许多特征值出现的问题中，极值特征值是需要计算的。同样，可以将hn的特征向量视为a的特征向量的近似值。

The Matlab function eigs is based on the computation of Ritz values. It computes the six eigenvalues of largest magnitude of a matrix A, and the call is [V, D] = eigs(A). More generally, to get the top k eigenvalues, use [V, D] = eigs(A, k).  
Matlab函数的特征值是基于Ritz值的计算。它计算矩阵A的最大数量的六个特征值，调用为[v，d]=eigs（a）。更一般地说，要得到顶部的k特征值，使用[v，d]=特征值（a，k）。

In the absence of rigorous theorems about error estimates, it is hard to make the above statements more precise; see Trefethen and Bau [171] (Lecture 34) for more on this subject.  
在缺乏关于误差估计的严格定理的情况下，很难使上述陈述更加精确；关于这一主题的更多信息，请参阅Trefethen和Bau[171]（第34课）。

However, if A is a symmetric (or Hermitian) matrix, then Hn is a symmetric (resp. Hermitian) tridiagonal matrix and more precise results can be shown; see Demmel [49] (Chapter 7, especially Section 7.2). We will consider the symmetric (and Hermitan) case in the next section, but first we show how Arnoldi iteration can be used to find approximations for the solution of a linear system Ax = b where A is invertible but of very large dimension m.  
但是，如果a是对称（或厄米特）矩阵，那么hn是对称（resp）。Hermitian）三对角矩阵和更精确的结果可以显示出来；见demmel[49]（第7章，特别是第7.2节）。我们将在下一节中考虑对称（和厄米坦）情况，但首先我们将展示如何使用Arnoldi迭代来寻找线性系统ax=b的近似解，其中a是可逆的，但m的尺寸非常大。

## 22.5 GMRES 22.5克

Suppose A is an invertible m×m matrix and let b be a nonzero vector in Cm. Let x0 = A−1b, the unique solution of Ax = b. It is not hard to show that x0 ∈ Kn(A,b) for some n ≤ m. In fact, there is a unique monic polynomial p(z) of minimal degree s ≤ m such that p(A)b = 0, so x0 ∈ Ks(A,b). Thus it makes sense to search for a solution of Ax = b in Krylov spaces of dimension m ≤ s. The idea is to find an approximation xn ∈ Kn(A,b) of x0 such that rn = b − Axn is minimized, that is, krnk2 = kb − Axnk2 is minimized over xn ∈ Kn(A,b).  
假设A是可逆M×M矩阵，B是非零向量，单位为厘米。设X0=a−1b，ax=b的唯一解，不难证明X0∈kn（a，b）对于一些n≤m，实际上存在一个极小阶s≤m的唯一Monic多项式p（z），使得p（a）b=0，所以X0∈ks（a，b）。因此，在维数m≤s的krylov空间中寻找ax=b的解是有意义的，其思想是求X0的近似值xn∈kn（a，b），使rn=b−axn最小化，即krnk2=kb−axnk2在xn∈kn（a，b）上最小化。

This minimization problem can be stated as  
这个最小化问题可以表述为

minimize krnk2 = kAxn − bk2 , xn ∈ Kn(A,b).  
最小化krnk2=kaxn−bk2，xn∈kn（a，b）。

This is a least-squares problem, and we know how to solve it (see Section 21.1). The quantity rn is known as the residual and the method which consists in minimizing krnk2 is known as GMRES, for generalized minimal residuals.  
这是一个最小二乘问题，我们知道如何解决它（见第21.1节）。量Rn被称为残差，对于广义最小残差，包含最小化krnk2的方法被称为gmres。

Now since (u1,...,un) is a basis of Kn(A,b) (since n ≤ s, no breakdown occurs, except for n = s), we may write xn = Uny, so our minimization problem is  
现在，因为（u1，…，un）是kn（a，b）的基础（因为n≤s，除了n=s，没有发生故障），我们可以写xn=uny，所以我们的最小化问题是

minimize kAUny − bk2 , y ∈ Cn.  
最小化kauny−bk2，y∈cn。

Since by (∗1) of Section 22.4, we have AUn = Un+1Hen, minimizing kAUny − bk2 is equivalent to minimizing kUn+1Heny − bk2 over Cm. Since Un+1Heny and b belong to the column space of Un+1, minimizing kUn+1Heny − bk2 is equivalent to minimizing .  
由于在第22.4节（1）中，我们得出了aun=un+1hen，最小化kauny−bk2等于在cm上最小化kun+1heny−bk2。由于un+1heny和b属于un+1的列空间，因此最小化kun+1heny−bk2等于最小化。

However, by construction,  
但是，通过施工，

,  
，

so our minimization problem can be stated as  
所以我们的最小化问题可以表述为

minimize kHeny − kbk2e1k2, y ∈ Cn.  
最小化kheny−kbk2e1k2，y∈cn。

The approximate solution of Ax = b is then  
ax=b的近似解是

xn = Uny.  
xn=uny。

Starting with u1 = b/kbk2 and with n = 1, the GMRES method runs n ≤ s Arnoldi iterations to find Un and Hen, and then runs a method to solve the least squares problem  
从u1=b/kbk2，n=1开始，GMRES方法进行n≤s阿诺尔底迭代，找到un和hen，然后运行一种求解最小二乘问题的方法。

minimize kHeny − kbk2e1k2, y ∈ Cn.  
最小化kheny−kbk2e1k2，y∈cn。

### 22.6. THE HERMITIAN CASE; LANCZOS ITERATION 22.6。赫米特案例；兰佐斯迭代

When krnk2 = kHeny−kbk2e1k2 is considered small enough, we stop and the approximate solution of Ax = b is then xn = Uny.  
当krnk2=kheny−kbk2e1k2足够小时，我们停止，ax=b的近似解为xn=uny。

There are ways of improving efficiency of the “naive” version of GMRES that we just presented; see Trefethen and Bau [171] (Lecture 35). We now consider the case where A is a Hermitian (or symmetric) matrix.  
我们刚刚介绍的GMRES“幼稚”版本有一些提高效率的方法；见Trefethen和Bau[171]（第35课）。我们现在考虑的情况是，a是一个厄米特（或对称）矩阵。

## 22.6 The Hermitian Case; Lanczos Iteration 22.6 Hermitian案例；Lanczos迭代

If A is an m×m symmetric or Hermitian matrix, then Arnoldi’s method is simpler and much more efficient. Indeed, in this case, it is easy to see that the upper Hessenberg matrices Hn are also symmetric (Hermitian respectively), and thus tridiagonal. Also, the eigenvalues of  
如果a是m×m对称矩阵或厄米特矩阵，那么阿诺迪方法简单而有效。事实上，在这种情况下，很容易看出上海森堡矩阵hn也是对称的（分别是赫米特矩阵），因此是三对角矩阵。另外，特征值

A and Hn are real. It is convenient to write  
a和hn是真的。写起来很方便

α1 β1   
α1β1\_

β1 α2 β2   
β1α2β2\_

Hn =  β2 α3 ... .  
hn=β2α3…。

   
 

 ... ... βn−1 βn−1 αn  
……βn−1βn−1αn

The recurrence (∗2) of Section 22.4 becomes the three-term recurrence  
第22.4节的复发（2）成为三期复发。

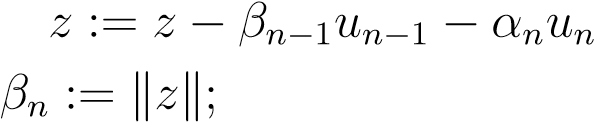
Aun = βn−1un−1 + αnun + βnun+1. (∗6)  
aun=βn−1un−1+αnun+βnun+1。（6）

We also have, so Arnoldi’s algorithm become the following algorithm known as Lanczos’ algorithm (or Lanczos iteration). The inner loop on j from 1 to n has been eliminated and replaced by a single assignment.  
我们也有，所以Arnoldi的算法变成了下面的算法，叫做Lanczos算法（或Lanczos迭代）。J上从1到n的内环已被消除，并被一个单独的赋值所取代。

Given an arbitrary nonzero vector b ∈ Cm, let u1 = b/kbk; for n = 1,2,3,... do  
给定任意非零向量b∈cm，设u1=b/kbk；对于n=1,2,3，…做

z := Aun;  
Z：=aun；

;  
；



if βn = 0 quit un+1 = z/βn  
如果βn=0，退出un+1=z/βn

When βn = 0, we say that we have a breakdown of the Lanczos iteration.  
当βn=0时，我们说我们有兰佐斯迭代的分解。

Versions of Proposition 22.5 and Theorem 22.6 apply to Lanczos iteration.  
命题22.5和定理22.6的版本适用于Lanczos迭代。

Besides being much more efficient than Arnoldi iteration, Lanczos iteration has the advantage that the Rayleigh–Ritz method for finding some of the eigenvalues of A as the eigenvalues of the symmetric (respectively Hermitian) tridiagonal matrix Hn applies, but there are more methods for finding the eigenvalues of symmetric (respectively Hermitian) tridiagonal matrices. Also theorems about error estimates exist. The version of Lanczos iteration given above may run into problems in floating point arithmetic. What happens is that the vectors uj may lose the property of being orthogonal, so it may be necessary to reorthogonalize them. For more on all this, see Demmel [49] (Chapter 7, in particular Section 7.2-7.4). The version of GMRES using Lanczos iteration is called MINRES.  
Lanczos迭代除了比Arnoldi迭代更有效外，还有一个优点，那就是瑞利-瑞兹方法用于寻找a的一些特征值，作为对称（分别是Hermitian）三对角矩阵hn的特征值，但是有更多的方法可以找到t。对称（分别是厄米特矩阵）三对角矩阵的特征值。还存在关于误差估计的定理。上面给出的Lanczos迭代版本可能在浮点运算中遇到问题。结果是，矢量UJ可能失去正交性，因此有必要对其进行重定位。有关更多信息，请参见demmel[49]（第7章，特别是第7.2-7.4节）。使用lanczos迭代的gmres版本称为minres。

We close our brief survey of methods for computing the eigenvalues and the eigenvectors of a matrix with a quick discussion of two methods known as power methods.  
我们结束了对计算矩阵特征值和特征向量的方法的简短调查，并对两种称为幂次法的方法进行了快速讨论。

## 22.7 Power Methods 22.7动力方法

Let A be an m × m complex or real matrix. There are two power methods, both of which yield one eigenvalue and one eigenvector associated with this vector:  
设A为m×m复矩阵或实矩阵。有两种功率方法，都会产生一个特征值和一个与此向量相关的特征向量：

1. Power iteration.  
   动力迭代。
2. Inverse (power) iteration.  
   逆（幂）迭代。

Power iteration only works if the matrix A has an eigenvalue λ of largest modulus, which means that if λ1,...,λm are the eigenvalues of A, then  
只有当矩阵A具有最大模的特征值λ时，方可进行幂次迭代，也就是说，如果λ1，…，λm是a的特征值，那么

|λ1| > |λ2| ≥ ··· ≥ |λm| ≥ 0.  
|λ1>λ2≥······≥λm≥0.

In particular, if A is a real matrix, then λ1 must be real (since otherwise there are two complex conjugate eigenvalues of the same largest modulus). If the above condition is satisfied, then power iteration yields λ1 and some eigenvector associated with it. The method is simple enough:  
特别是，如果a是一个实矩阵，那么λ1必须是实矩阵（否则有两个相同最大模的复共轭特征值）。如果满足上述条件，则功率迭代得到λ1及其相关的一些特征向量。方法非常简单：

Pick some initial unit vector x0 and compute the following sequence (xk), where  
选取一些初始单位向量X0并计算以下序列（XK），其中

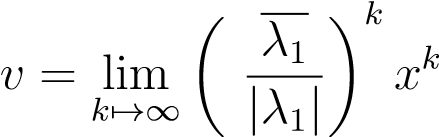
.  
.

We would expect that (xk) converges to an eigenvector associated with λ1, but this is not quite correct. The following results are proven in Serre [151] (Section 13.5.1). First assume that λ1 = 0.6  
我们期望（xk）收敛到与λ1相关的特征向量，但这并不完全正确。以下结果在SERRE[151]中得到证实（第13.5.1节）。首先假设λ1=0.6

We have  
我们有

.  
.

If A is a complex matrix which has a unique complex eigenvalue λ1 of largest modulus, then  
如果A是一个具有最大模的唯一复特征值λ1的复矩阵，那么



### 22.7. POWER METHODS 22.7。功率法

is a unit eigenvector of A associated with λ1. If λ1 is real, then  
是与λ1相关联的a的单位特征向量。如果λ1为真，则

v = lim xk k7→∞  
V=lim xk k7→∞

is a unit eigenvector of A associated with λ1. Actually some condition on x0 is needed: x0 must have a nonzero component in the eigenspace E associated with λ1 (in any direct sum of Cm in which E is a summand).  
是与λ1相关联的a的单位特征向量。实际上，在X0上需要一些条件：X0必须在与λ1相关的特征空间e中有一个非零分量（在任何直接和cm中，e是和）。

The eigenvalue λ1 is found as follows. If λ1 is complex, and if vj = 06 is any nonzero coordinate of v, then  
特征值λ1如下所示。如果λ1是复数，并且vj=06是v的任何非零坐标，那么

.  
.

If λ1 is real, then we can define the sequence (λ(k)) by  
如果λ1是实的，那么我们可以通过以下方式定义序列（λ（k））：

λ(k+1) = (xk+1)∗Axk+1, k ≥ 0,  
λ（k+1）=（xk+1）axk+1，k≥0，

and we have  
我们有

λ1 = lim λ(k). k7→∞  
λ1=limλ（k）。K7→∞

Indeed, in this case, since v = limk7→∞ xk and v is a unit eigenvector for λ1, we have  
实际上，在这种情况下，由于v=limk7→∞xk和v是λ1的单位特征向量，我们得到

lim λ(k) = lim (xk+1)∗Axk+1 = v∗Av = λ1v∗v = λ1. k7→∞ k7→∞  
limλ（k）=lim（xk+1）axk+1=v av=λ1v v=λ1.K7→∞K7→∞

Note that since xk+1 is a unit vector, (xk+1)∗Axk+1 is a Rayleigh ratio.  
注意，由于xk+1是单位向量，（xk+1）axk+1是瑞利比。

If A is a Hermitian matrix, then the eigenvalues are real and we can say more about the rate of convergence, which is not great (only linear). For details, see Trefethen and Bau [171] (Lecture 27).  
如果a是一个厄米矩阵，那么特征值是真实的，我们可以说更多的收敛速度，这不是很大（只有线性）。有关详细信息，请参阅Trefetten和Bau[171]（第27课）。

If λ1 = 0, then there is some power ` < m such that Ax` = 0.  
如果λ1=0，则存在一些功率`<m，使得ax`=0。

The inverse iteration method is designed to find an eigenvector associated with an eigenvalue λ of A for which we know a good approximation µ.  
反迭代法的目的是找到一个特征向量与一个我们知道一个很好的近似值μ的特征值λ相关。

Pick some initial unit vector x0 and compute the following sequences (wk) and (xk), where wk+1 is the solution of the system  
选取一些初始单位向量X0，计算以下序列（wk）和（xk），其中wk+1是系统的解

(A − µI)wk+1 = xk equivalently wk+1 = (A − µI)−1xk, k ≥ 0,  
（a−μi）wk+1=xk等于wk+1=（a−μi）−1xk，k≥0，

and  
和

.  
.

The following result is proven in Ciarlet [41] (Theorem 6.4.1).  
以下结果在Ciarlet[41]中得到证明（定理6.4.1）。

Proposition 22.7. Let A be an m × m diagonalizable (complex or real) matrix with eigenvalues λ1,...,λm, and let λ = λ` be an arbitrary eigenvalue of A (not necessary simple).  
提案22.7.设a为特征值为λ1，…，λm的m×m可对角化（复或实）矩阵，并设λ=λ`为a的任意特征值（不必简单）。

For any µ such that  
对于任何这样的

µ =6 λ and |µ − λ| < |µ − λj| for all j =6 `,  
μ=6λ和μ−λ<μ−λj对于所有j=6`，

if x0 does not belong to the subspace spanned by the eigenvectors associated with the eigenvalues λj with j =6 `, then  
如果X0不属于特征值λj（j=6`）相关特征向量所跨越的子空间，则

−  
-

where v is an eigenvector associated with λ. Furthermore, if both λ and µ are real, we have  
其中v是与λ相关的特征向量。此外，如果λ和μ都是真的，我们有

|  |  |
| --- | --- |
| lim xk = v k7→∞ 网络错误  lim (−1)kxk = v 网络错误 | if µ < λ, 网络错误  if µ > λ. 网络错误 |

k7→∞  
K7→∞

Also, if we define the sequence (λ(k)) by  
另外，如果我们用

λ(k+1) = (xk+1)∗Axk+1,  
λ（k+1）=（xk+1）axk+1，

then  
然后

lim λ(k+1) = λ.  
limλ（k+1）=λ。

k7→∞  
K7→∞

The condition of x0 may seem quite stringent, but in practice, a vector x0 chosen at random usually satisfies it.  
X0的条件可能看起来相当严格，但在实践中，随机选择的向量X0通常满足它。

If A is a Hermitian matrix, then we can say more. In particular, the inverse iteration algorithm can be modified to make use of the newly computed λ(k+1) instead of µ, and an even faster convergence is achieved. Such a method is called the Rayleigh quotient iteration. When it converges (which is for almost all x0), this method eventually achieves cubic convergence, which is remarkable. Essentially, this means that the number of correct digits is tripled at every iteration. For more details, see Trefethen and Bau [171] (Lecture 27) and Demmel [49] (Section 5.3.2).  
如果A是一个厄米矩阵，那么我们可以说更多。特别是，可以修改逆迭代算法，以利用新计算的λ（k+1）而不是μ，从而实现更快的收敛。这种方法叫做瑞利商迭代。当它收敛时（几乎是所有的X0），这种方法最终达到了三次收敛，这是显著的。从本质上来说，这意味着在每次迭代中正确数字的数量是三倍。有关更多详细信息，请参阅Trefetten和Bau[171]（第27讲）和Demmel[49]（第5.3.2节）。

## 22.8 Summary 22.8总结

The main concepts and results of this chapter are listed below:  
本章的主要概念和结果如下：

* QR iteration, QR algorithm.  
  二维码迭代，二维码算法。
* Upper Hessenberg matrices.  
  上海森堡矩阵。
* Householder matrix.  
  户主矩阵。

### 22.9. PROBLEMS 22.9。问题

* Unreduced and reduced Hessenberg matrices.  
  未简化和约化的Hessenberg矩阵。
* Deflation.  
  通货紧缩。
* Shift.  
  换档。
* Wilkinson shift.  
  威尔金森轮班。
* Double shift.  
  双班制。
* Francis shift.  
  弗朗西斯变换。
* Implicit shifting.  
  隐性转变。
* Implicit Q-theorem.  
  隐式Q定理。
* Arnoldi iteration.  
  阿诺迪迭代。
* Breakdown of Arnoldi iteration.  
  阿诺迪迭代的分解。
* Krylov subspace.  
  Krylov子空间。
* Rayleigh–Ritz method.  
  瑞利-里兹法。
* Ritz values, Arnoldi estimates.  
  里兹值，阿诺迪估计。
* Residual.  
  剩余。
* GMRES  
  GMRES
* Lanczos iteration.  
  Lanczos迭代。
* Power iteration.  
  动力迭代。
* Inverse power iteration.  
  逆功率迭代。
* Rayleigh ratio.  
  瑞利比。

## 22.9 Problems 22.9问题

Problem 22.1. Prove Theorem 22.2; see Problem 12.7.  
问题22.1。证明定理22.2；见问题12.7。

Problem 22.2. Prove that if a matrix A is Hermitian (or real symmetric), then any Hessenberg matrix H similar to A is Hermitian tridiagonal (real symmetric tridiagonal).  
问题22.2。证明了如果矩阵A是厄米特矩阵（或实对称），那么任何与A相似的海森堡矩阵H都是厄米特三对角矩阵（实对称三对角）。

Problem 22.3. For any matrix (real or complex) A, if A = QR is a QR-decomposition of A using Householder reflections, prove that if A is upper Hessenberg then so is Q.  
问题22.3。对于任何矩阵（实矩阵或复矩阵）a，如果a=qr是使用户主反射的qr分解，证明如果a是上赫斯伯格，那么q也是。

Problem 22.4. Prove that if A is upper Hessenberg, then the matrices Ak obtained by applying the QR-algorithm are also upper Hessenberg.  
问题22.4.证明如果a是上海森堡，那么应用qr算法得到的矩阵ak也是上海森堡。

Problem 22.5. Prove the implicit Q theorem. This theorem says that if A is reduced to upper Hessenberg form as A = UHU∗ and if H is unreduced (hi+1i = 06 for i = 1,...,n−1), then the columns of index 2,...,n of U are determined by the first column of U up to sign;  
问题22.5。证明隐式Q定理。这个定理表明，如果a被简化为上赫森伯格形式a=u h u，如果h不被简化（i=1，…，n−1时hi+1i=06），那么索引2，…，n的列由u的第一列到符号决定；

Problem 22.6. Read Section 7.5 of Golub and Van Loan [80] and implement their version of the QR-algorithm with shifts.  
问题22.6.阅读Golub和van Loan[80]的第7.5节，并通过移位实现他们的QR算法版本。

Problem 22.7. If an Arnoldi iteration has a breakdown at stage n, that is, hn+1 = 0, then we found the first unreduced block of the Hessenberg matrix H. Prove that the eigenvalues of Hn are eigenvalues of A.  
问题22.7。如果阿诺迪迭代在n阶段有一个分解，即hn+1=0，那么我们就找到了Hessenberg矩阵h的第一个不可约块，证明hn的特征值是a的特征值。

Problem 22.8. Prove Theorem 22.6.  
问题22.8。证明定理22.6。

Problem 22.9. Implement GRMES and test it on some linear systems.  
问题22.9.实现GRMES并在一些线性系统上进行测试。

Problem 22.10. State and prove versions of Proposition 22.5 and Theorem 22.6 for the Lanczos iteration.  
问题22.10。陈述并证明关于Lanczos迭代的22.5号命题和22.6号定理的版本。

Problem 22.11. Prove the results about the power iteration method stated in Section 22.7.  
问题22.11。证明第22.7节所述的功率迭代法的结果。

Problem 22.12. Prove the results about the inverse power iteration method stated in Section 22.7.  
问题22.12。证明第22.7节所述逆功率迭代法的结果。

Problem 22.13. Implement and test the power iteration method and the inverse power iteration method.  
问题22.13。实现并测试了功率迭代法和逆功率迭代法。

Problem 22.14. Read Lecture 27 in Trefethen and Bau [171] and implement and test the Rayleigh quotient iteration method.  
问题22.14。阅读Trefethen和Bau[171]中的第27讲，实现并测试瑞利商迭代法。

Part II  
第二部分

Affine and Projective Geometry  
仿射几何和射影几何

693  
六百九十三

Chapter 23  
第二十三章

# Basics of Affine Geometry 仿射几何基础

L’alg`ebre n’est qu’une g´eom´etrie ´ecrite; la g´eom´etrie n’est qu’une alg`ebre figur´ee.  
l'alg'ebre n'est qu'une g'eom'etrie'ecrite；la g'eom'etrie n'est qu'une alg'ebre figur ee'ee。

—Sophie Germain  
-索菲日尔曼

## 23.1 Affine Spaces 23.1仿射空间

Geometrically, curves and surfaces are usually considered to be sets of points with some special properties, living in a space consisting of “points.” Typically, one is also interested in geometric properties invariant under certain transformations, for example, translations, rotations, projections, etc. One could model the space of points as a vector space, but this is not very satisfactory for a number of reasons. One reason is that the point corresponding to the zero vector (0), called the origin, plays a special role, when there is really no reason to have a privileged origin. Another reason is that certain notions, such as parallelism, are handled in an awkward manner. But the deeper reason is that vector spaces and affine spaces really have different geometries. The geometric properties of a vector space are invariant under the group of bijective linear maps, whereas the geometric properties of an affine space are invariant under the group of bijective affine maps, and these two groups are not isomorphic. Roughly speaking, there are more affine maps than linear maps.  
在几何上，曲线和曲面通常被认为是一组具有某些特殊性质的点，它们生活在一个由“点”组成的空间中。通常，人们也对在某些变换下不变的几何性质感兴趣，例如，平移、旋转、投影离子等。人们可以将点的空间建模为矢量空间，但由于许多原因，这并不十分令人满意。一个原因是与零向量（0）相对应的点（称为原点）在没有理由拥有特权原点的情况下起着特殊的作用。另一个原因是某些概念，如并行性，处理起来很尴尬。但深层次的原因是向量空间和仿射空间的几何性质是不同的。向量空间的几何性质在双射线性映射群下是不变的，而仿射空间的几何性质在双射仿射映射群下是不变的，这两个群不是同构的。大致来说，仿射映射比线性映射多。

Affine spaces provide a better framework for doing geometry. In particular, it is possible to deal with points, curves, surfaces, etc., in an intrinsic manner, that is, independently of any specific choice of a coordinate system. As in physics, this is highly desirable to really understand what is going on. Of course, coordinate systems have to be chosen to finally carry out computations, but one should learn to resist the temptation to resort to coordinate systems until it is really necessary.  
仿射空间为几何提供了一个更好的框架。特别是，可以以一种内在的方式处理点、曲线、曲面等，即独立于坐标系的任何特定选择。在物理学中，真正理解正在发生的事情是非常需要的。当然，为了最终进行计算，必须选择坐标系，但在真正必要之前，人们应该学会抵制使用坐标系的诱惑。

Affine spaces are the right framework for dealing with motions, trajectories, and physical forces, among other things. Thus, affine geometry is crucial to a clean presentation of kinematics, dynamics, and other parts of physics (for example, elasticity). After all, a rigid motion is an affine map, but not a linear map in general. Also, given an m × n matrix A  
仿射空间是处理运动、轨迹和物理力等问题的正确框架。因此，仿射几何对于运动学、动力学和其他物理部分（例如弹性）的清晰呈现至关重要。毕竟，刚性运动是仿射映射，而不是一般的线性映射。另外，给定m×n矩阵a

695  
六百九十五

and a vector b ∈ Rm, the set U = {x ∈ Rn | Ax = b} of solutions of the system Ax = b is an affine space, but not a vector space (linear space) in general.  
而向量b∈rm，系统解a x=b的集u=x∈rn ax=b是仿射空间，但一般不是向量空间（线性空间）。

Use coordinate systems only when needed!  
仅在需要时使用坐标系！

This chapter proceeds as follows. We take advantage of the fact that almost every affine concept is the counterpart of some concept in linear algebra. We begin by defining affine spaces, stressing the physical interpretation of the definition in terms of points (particles) and vectors (forces). Corresponding to linear combinations of vectors, we define affine combinations of points (barycenters), realizing that we are forced to restrict our attention to families of scalars adding up to 1. Corresponding to linear subspaces, we introduce affine subspaces as subsets closed under affine combinations. Then, we characterize affine subspaces in terms of certain vector spaces called their directions. This allows us to define a clean notion of parallelism. Next, corresponding to linear independence and bases, we define affine independence and affine frames. We also define convexity. Corresponding to linear maps, we define affine maps as maps preserving affine combinations. We show that every affine map is completely defined by the image of one point and a linear map. Then, we investigate briefly some simple affine maps, the translations and the central dilatations. At this point, we give a glimpse of affine geometry. We prove the theorems of Thales, Pappus, and Desargues. After this, the definition of affine hyperplanes in terms of affine forms is reviewed. The section ends with a closer look at the intersection of affine subspaces.  
本章内容如下。我们利用了这样一个事实：几乎每一个仿射概念都是线性代数中某些概念的对应物。我们首先定义仿射空间，强调用点（粒子）和向量（力）来解释定义。对应于向量的线性组合，我们定义了点的仿射组合（重心），意识到我们被迫将注意力限制在最多1个标量的族上。针对线性子空间，引入仿射子空间作为仿射组合下的闭子集。然后，我们用称为方向的向量空间来描述仿射子空间。这允许我们定义一个清晰的并行性概念。其次，对应于线性独立和基，定义了仿射独立和仿射框架。我们也定义了凸性。对应于线性映射，我们将仿射映射定义为保留仿射组合的映射。我们证明了每一个仿射映射都完全由一个点的图像和一个线性映射定义。然后，我们简单地研究了一些简单的仿射映射、翻译和中心扩张。在这一点上，我们给出了仿射几何的一瞥。我们证明了泰雷兹、帕普斯和德沙格的定理。在此基础上，回顾了仿射超平面在仿射形式上的定义。该部分以更仔细的观察仿射子空间的交叉点结束。

Our presentation of affine geometry is far from being comprehensive, and it is biased toward the algorithmic geometry of curves and surfaces. For more details, the reader is referred to Pedoe [132], Snapper and Troyer [157], Berger [11, 12], Coxeter [44], Samuel [138], Tisseron [170], Fresnel [66], Vienne [179], and Hilbert and Cohn-Vossen [90].  
我们对仿射几何的描述还远远不够全面，它偏向于曲线和曲面的算法几何。关于更多细节，读者可以参考Pedoe[132]、Snapper和Troyer[157]、Berger[11，12]、Coxeter[44]、Samuel[138]、Tisseron[170]、Fresnel[66]、Vienne[179]和Hilbert和Cohn Vossen[90]。

Suppose we have a particle moving in 3D space and that we want to describe the trajectory of this particle. If one looks up a good textbook on dynamics, such as Greenwood [82], one finds out that the particle is modeled as a point, and that the position of this point x is determined with respect to a “frame” in R3 by a vector. Curiously, the notion of a frame is rarely defined precisely, but it is easy to infer that a frame is a pair (O,(e1,e2,e3)) consisting of an origin O (which is a point) together with a basis of three vectors (e1,e2,e3). For example, the standard frame in R3 has origin O = (0,0,0) and the basis of three vectors e1 = (1,0,0), e2 = (0,1,0), and e3 = (0,0,1). The position of a point x is then defined by the “unique vector” from O to x.  
假设有一个粒子在三维空间中运动，我们想描述这个粒子的轨迹。如果你查阅了一本很好的动力学教科书，比如Greenwood[82]，你会发现粒子被建模为一个点，这个点x的位置是由一个向量相对于r3中的“帧”来确定的。奇怪的是，帧的概念很少被精确定义，但很容易推断出帧是一对（o，（e1，e2，e3）），由原点o（即点）和三个矢量（e1，e2，e3）的基组成。例如，r3中的标准帧具有原点o=（0,0,0）和三个向量e1=（1,0,0）、e2=（0,1,0）和e3=（0,0,1）。点X的位置由从O到X的“唯一向量”定义。

But wait a minute, this definition seems to be defining frames and the position of a point without defining what a point is! Well, let us identify points with elements of R3. If so, given any two points a = (a1,a2,a3) and b = (b1,b2,b3), there is a unique free vector, denoted by →−ab, from a to b, the vector →−ab = (b1 − a1,b2 − a2,b3 − a3). Note that  
但是等一下，这个定义似乎定义了框架和点的位置，而没有定义点是什么！好吧，让我们用r3的元素来确定点。如果是这样，在任意两点a=（a1、a2、a3）和b=（b1、b2、b3）下，有一个唯一的自由矢量，用→−ab表示，从a到b，矢量→−ab=（b1−a1、b2−a2、b3−a3）。注意

b = a + →−ab,  
B=A+→AB，

*O*

*a*

*b*

−

→

*ab*

Figure 23.1: Points and free vectors.  
图23.1：点和自由矢量。

addition being understood as addition in R3. Then, in the standard frame, given a point x = (x1,x2,x3), the position of x is the vector Ox−→ = (x1,x2,x3), which coincides with the point itself. In the standard frame, points and vectors are identified. Points and free vectors are illustrated in Figure 23.1.  
加成被理解为r3中的加成。然后，在标准帧中，给定一个点x=（x1，x2，x3），x的位置是向量ox−→=（x1，x2，x3），它与点本身重合。在标准框架中，识别点和向量。点和自由矢量如图23.1所示。

What if we pick a frame with a different origin, say Ω = (ω1,ω2,ω3), but the same basis vectors (e1,e2,e3)? This time, the point x = (x1,x2,x3) is defined by two position vectors:  
如果我们选取一个原点不同的帧，比如Ω=（ω1，ω2，ω3），但基向量相同（e1，e2，e3），会怎么样？这一次，点x=（x1，x2，x3）由两个位置矢量定义：

Ox−→ = (x1,x2,x3)  
Ox−→=（x1，x2，x3）

in the frame (O,(e1,e2,e3)) and  
在帧（o，（e1，e2，e3））和

Ω−→x = (x1 − ω1,x2 − ω2,x3 − ω3)  
Ω−→x=（x1−ω1，x2−ω2，x3−ω3）

in the frame (Ω,(e1,e2,e3)). See Figure 23.2.  
在框架中（Ω，（e1，e2，e3））。见图23.2。

This is because and O−→Ω = (ω1,ω2,ω3).  
这是因为和O−→Ω=（ω1，ω2，ω3）。

We note that in the second frame (Ω,(e1,e2,e3)), points and position vectors are no longer identified. This gives us evidence that points are not vectors. It may be computationally convenient to deal with points using position vectors, but such a treatment is not frame invariant, which has undesirable effects.  
我们注意到，在第二帧（Ω，（e1，e2，e3））中，不再识别点和位置矢量。这给了我们证据，证明点不是向量。使用位置向量处理点可能在计算上很方便，但这种处理不是帧不变的，这会产生不良的效果。

Inspired by physics, we deem it important to define points and properties of points that are frame invariant. An undesirable side effect of the present approach shows up if we attempt to define linear combinations of points. First, let us review the notion of linear combination of vectors. Given two vectors u and v of coordinates (u1,u2,u3) and (v1,v2,v3) with respect  
受物理学的启发，我们认为定义帧不变的点和点的性质很重要。如果我们试图定义点的线性组合，则会出现当前方法的不良副作用。首先，让我们回顾向量线性组合的概念。给定坐标（U1、U2、U3）和（V1、V2、V3）的两个向量u和v

*Ω*

x

3

e

3

e

2

e

2

e

1

e

1

e

*Ω*

O

x

x

O

Figure 23.2: The two position vectors for the point x.  
图23.2:X点的两个位置矢量。

to the basis (e1,e2,e3), for any two scalars λ,µ, we can define the linear combination λu+µv as the vector of coordinates  
对于基（e1，e2，e3），对于任意两个标量λ，μ，我们可以定义线性组合λu+μv作为坐标矢量。

(λu1 + µv1,λu2 + µv2,λu3 + µv3).  
（λu1+μv1，λu2+μv2，λu3+μv3）。

If we choose a different basis () and if the matrix P expressing the vectors ( over the basis (e1,e2,e3) is  
如果我们选择一个不同的基（），并且表示向量的矩阵p（相对于基（e1，e2，e3）是

,  
，

which means that the columns of P are the coordinates of the e0j over the basis (e1,e2,e3), since  
这意味着p列是e0j在基上的坐标（e1，e2，e3），因为



and  
和

,  
，

it is easy to see that the coordinates (u1,u2,u3) and (v1,v2,v3) of u and v with respect to the basis (e1,e2,e3) are given in terms of the coordinates () and (and v with respect to the basis () by the matrix equations  
很容易看出，u和v相对于基（e1，e2，e3）的坐标（u1，u2，u3）和（v1，v2，v3）由矩阵方程给出，相对于基（e1，e2，e3）的坐标（）和（和v相对于基（）。

and .  
而且。

From the above, we get  
从上面我们可以看到

and ,  
而且，

and by linearity, the coordinates  
通过线性，坐标



of λu + µv with respect to the basis () are given by  
关于基础（）的λu+μv的公式如下：

.  
.

Everything worked out because the change of basis does not involve a change of origin. On the other hand, if we consider the change of frame from the frame (O,(e1,e2,e3)) to the frame (Ω,(e1,e2,e3)), where O−→Ω = (ω1,ω2,ω3), given two points a, b of coordinates (a1,a2,a3) and (b1,b2,b3) with respect to the frame (O,(e1,e2,e3)) and of coordinates (  
一切都是因为基础的改变并不涉及到原产地的改变。另一方面，如果我们考虑从帧（o，（e1，e2，e3））到帧（Ω，（e1，e2，e3））的变化，其中o−→（

) with respect to the frame (Ω,(e1,e2,e3)), since  
）关于框架（Ω，（e1，e2，e3）），因为

(a01,a02,a03) = (a1 − ω1,a2 − ω2,a3 − ω3)  
（a01，a02，a03）=（a1−ω1，a2−ω2，a3−ω3）

and  
和

(b01,b02,b03) = (b1 − ω1,b2 − ω2,b3 − ω3),  
（b01、b02、b03）=（b1−ω1、b2−ω2、b3−ω3）

the coordinates of λa + µb with respect to the frame (O,(e1,e2,e3)) are  
λa+μb相对于框架（o，（e1，e2，e3））的坐标为

(λa1 + µb1,λa2 + µb2,λa3 + µb3),  
（λa1+μb1，λa2+μb2，λa3+μb3）、

but the coordinates of λa + µb with respect to the frame (Ω,(e1,e2,e3)) are  
但是，λa+μb相对于框架的坐标（Ω，（e1，e2，e3））是

(λa1 + µb1 − (λ + µ)ω1,λa2 + µb2 − (λ + µ)ω2,λa3 + µb3 − (λ + µ)ω3),  
（λa1+μb1−（λ+μ）ω1，λa2+μb2−（λ+μ）ω2，λa3+μb3−（λ+μ）ω3）

which are different from  
有什么不同

(λa1 + µb1 − ω1,λa2 + µb2 − ω2,λa3 + µb3 − ω3),  
（λa1+μb1-ω1，λa2+μb2-ω2，λa3+μb3-ω3）

unless λ + µ = 1. See Figure 23.3.  
除非λ+μ=1。见图23.3。

Thus, we have discovered a major difference between vectors and points: The notion of linear combination of vectors is basis independent, but the notion of linear combination of points is frame dependent. In order to salvage the notion of linear combination of points, some restriction is needed: The scalar coefficients must add up to 1.  
因此，我们发现了向量和点之间的一个主要区别：向量的线性组合的概念与基无关，但点的线性组合的概念与帧相关。为了恢复点的线性组合的概念，需要一些限制：标量系数必须加起来为1。

3

e

3

e

2

e

2

e

1

e

1

e

*Ω*

O

=

(3,4,

5)

a = (1,1,1)

b = (2,3,1)

a + b = (3,4,2)

=

(0, 0,

-3)

3

e

3

e

2

e

2

e

1

e

1

e

*Ω*

O

=

(3,4,

5)

a = (-2,-3,-4)

b = (-1,-1,-4)

a + b

=

(-3, -4,

-8)

=

(0, 0,

-3)

Figure 23.3: The top figure shows the location of the “point” sum a + b with respect to the frame (O,(e1,e2,e3)), while the bottom figure shows the location of the “point” sum a + b with respect to the frame (Ω,(e1,e2,e3)).  
图23.3：上图显示了“点”和A+B相对于帧（o，（e1，e2，e3））的位置，而下图显示了“点”和A+B相对于帧（Ω，（e1，e2，e3））的位置。

A clean way to handle the problem of frame invariance and to deal with points in a more intrinsic manner is to make a clearer distinction between points and vectors. We duplicate R3 into two copies, the first copy corresponding to points, where we forget the vector space structure, and the second copy corresponding to free vectors, where the vector space structure is important. Furthermore, we make explicit the important fact that the vector space R3 acts on the set of points R3 : Given any point a = (a1,a2,a3) and any vector v = (v1,v2,v3), we obtain the point  
处理帧不变性问题和处理更内在的点的一个干净方法是对点和向量进行更清晰的区分。我们将r3复制成两个副本，第一个副本对应于点，忽略了向量空间结构，第二个副本对应于自由向量，其中向量空间结构很重要。此外，我们明确了向量空间r3作用于点r3集合的重要事实：给定任意点a=（a1，a2，a3）和任意向量v=（v1，v2，v3），我们得到该点。

a + v = (a1 + v1,a2 + v2,a3 + v3),  
A+V=（A1+V1，A2+V2，A3+V3）

which can be thought of as the result of translating a to b using the vector v. We can imagine that v is placed such that its origin coincides with a and that its tip coincides with b. This action +: R3 × R3 → R3 satisfies some crucial properties. For example,  
这可以被认为是用向量v将a转换为b的结果。我们可以想象，v的位置使其原点与a重合，其尖端与b重合。这个作用+：r3×r3→r3满足一些关键特性。例如，

a + 0 = a,  
A+0=A，

(a + u) + v = a + (u + v),  
（A+U）+V=A+（U+V）

and for any two points a,b, there is a unique free vector →−ab such that  
对于任意两点a，b，有一个独特的自由矢量→−ab，这样

b = a + →−ab.  
B=A+→AB。

It turns out that the above properties, although trivial in the case of R3, are all that is needed to define the abstract notion of affine space (or affine structure). The basic idea is to consider two (distinct) sets E and →−E, where E is a set of points (with no structure) and →−E is a vector space (of free vectors) acting on the set E.  
结果表明，上述属性虽然在R3的情况下微不足道，但却是定义仿射空间（或仿射结构）抽象概念所需的全部属性。基本思想是考虑两个不同的集合e和→−e，其中e是一组点（没有结构），而→−e是作用于集合e的向量空间（自由向量）。

Did you say “A fine space”?  
你说“好地方”了吗？

Intuitively, we can think of the elements of →−E as forces moving the points in E, considered as physical particles. The effect of applying a force (free vector) u ∈ →−E to a point a ∈ E is a translation. By this, we mean that for every force u ∈ →−E, the action of the force u is to “move” every point a ∈ E to the point a+u ∈ E obtained by the translation corresponding→− to u viewed as a vector. Since translations can be composed, it is natural that E is a vector space.  
直观地说，我们可以把→−e的元素看作是移动e中点的力，被认为是物理粒子。将力（自由矢量）u∈→−e施加到点a∈e上的效果是平移。由此，我们的意思是，对于每一个力u∈→−e，力u的作用是“移动”每一个点a∈e到点a+u∈e，这是由对应的→−到u的平移得到的，被视为一个向量。因为翻译可以组合，所以e是向量空间是很自然的。

For simplicity, it is assumed that all vector spaces under consideration are defined over the field R of real numbers. Most of the definitions and results also hold for an arbitrary field K, although some care is needed when dealing with fields of characteristic different from zero. It is also assumed that all families (λi)i∈I of scalars have finite support. Recall that a family (λi)i∈I of scalars has finite support if λi = 0 for all i ∈ I −J, where J is a finite subset of I. Obviously, finite families of scalars have finite support, and for simplicity, the reader may assume that all families of scalars are finite. The formal definition of an affine space is as follows.  
为了简单起见，假设所有考虑的向量空间都是在实数的r域上定义的。大多数定义和结果也适用于任意场k，尽管在处理与零不同的特征场时需要注意一些。并假定所有的标量族（λi）i∈i都有有限的支持。回想一个标量的族（λi）i∈i具有有限的支持，如果λi=0表示所有i∈i−j，其中j是i的有限子集。显然，标量的有限族具有有限的支持，为了简单起见，读者可以假定所有标量的族都是有限的。仿射空间的形式定义如下。

Definition 23.1. An affine space is either the degenerate space reduced to the empty set, or a triple consisting of a nonempty set (of points), a vector space →−E (of translations, or free vectors), and an action +: E × E → E, satisfying the following conditions.  
定义23.1.仿射空间可以是退化空间降为空集，也可以是由非空集（点）、向量空间→−e（平移或自由向量）和操作+：e×e→e组成的三重空间，满足以下条件。

(A1) a + 0 = a, for every a ∈ E.  
（a1）a+0=a，每a∈e。

(A2) (a + u) + v = a + (u + v), for every a ∈ E, and every u,v ∈ →−E.  
（a2）（a+u）+v=a+（u+v），对于每一个a∈e，以及每一个u，v∈→−e。

(A3) For any two points a,b ∈ E, there is a uniquesuch that a + u = b.  
（a3）对于任意两点a，b∈e，存在一个唯一性，使得a+u=b。

The unique vector u ∈ →−E such that a + u = b is denoted by →−ab, or sometimes by ab, or even by b − a. Thus, we also write →−  
唯一向量u∈→−e，这样a+u=b用→−ab表示，有时用ab表示，甚至用b−a表示。因此，我们也写→−

b = a + ab  
B=A+AB

(or b = a + ab, or even b = a + (b − a)). →− →−  
（或b=a+ab，甚至b=a+（b−a））。→–→–

The dimension of the affine space is the dimension dim(E) of the vector space →−E. For simplicity, it is denoted by dim(E).  
仿射空间的维数是向量空间的维数dim（e）--e。为了简单起见，它用dim（e）表示。

*E*

−

→

*E*

*a*

*b*

=

*a*

+

*u*

*c*

=

*a*

+

*w*

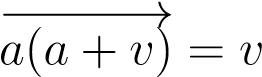
*u*

*v*

*w*

Figure 23.4: Intuitive picture of an affine space.  
图23.4：仿射空间的直观图片。

Conditions (A1) and (A2) say that the (abelian) group →−E acts on E, and Condition (A3) says that →−E acts transitively and faithfully on E. Note that  
条件（a1）和（a2）表示（abelian）组→−e对e起作用，条件（a3）表示→−e对e起过渡和忠实的作用。注意：



for all a ∈ E and all v ∈ →−E, since ) is the unique vector such that  
对于所有a∈e和所有v∈→−e，因为）是唯一的向量，这样

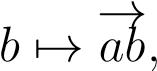
Thus, b = a + v is equivalent to ab = v. Figure 23.4 gives an intuitive picture of an affine space. It is natural to think of all vectors as having the same origin, the null vector.  
因此，b=a+v等于ab=v。图23.4给出了仿射空间的直观图像。很自然地，所有向量都有相同的原点，即零向量。

The axioms defining an affine space can be interpreted intuitively as saying that E and →−E are two different ways of looking at the same object, but wearing different sets of glasses, the second set of glasses depending on the choice of an “origin” in E. Indeed, we can choose to look at the points in E, forgetting that every pair (a,b) of points defines a unique vector →−ab in →−E, or we can choose to look at the vectors u in →−E, forgetting the points in E. Furthermore, if we also pick any point a in E, a point that can be viewed as an origin in E, then we can recover all the points in E as the translated points a + u for all u ∈ →−E. This can be formalized by defining two maps between E and →−E.  
定义仿射空间的公理可以直观地解释为，e和→−e是观察同一物体的两种不同方式，但戴着不同的眼镜，第二组眼镜取决于e中“原点”的选择。实际上，我们可以选择观察t。他指向e，忘记了每对（a，b）点定义了一个唯一的向量→–––––––––––––––––––––––––––––––––––––––––––––––––––––––––e中的所有点作为所有u∈→−e的转换点a+u。这可以通过定义e和→−e之间的两个映射形式化。

For every a ∈ E, consider the mapping from →−E to E given by  
对于每一个a∈e，考虑由

u 7→ a + u,  
U 7→A+U，

where u ∈ →−E, and consider the mapping from E to →−E given by  
式中u∈→−e，并考虑e到→−e的映射，由



where b ∈ E. The composition of the first mapping with the second is  
式中b∈e。第一个映射与第二个映射的组合是

,  
，

which, in view of (A3), yields u. The composition of the second with the first mapping is b 7→ →−ab 7→ a + →−ab,  
根据（a3），得出u。第二个具有第一个映射的组合是b 7→→−ab7→a+→−a b，

which, in view of (A3), yields b. Thus, these compositions are the identity from →−E to →−E and the identity from E to E, and the mappings are both bijections.  
从（a3）的角度来看，生成b。因此，这些组成是从→−e到→−e的同一性和从e到e的同一性，并且映射都是双射。

When we identify E with →−E via the mapping b →7 →−ab, we say that we consider E as the vector space obtained by taking a as the origin in E, and we denote it by Ea. Because Ea is a vector space, to be consistent with our notational conventions we should use the notation (using an arrow), instead of Ea. However, for simplicity, we stick to the notation Ea.  
当我们通过映射b→7→ab将e与→−e标识时，我们认为e是以a作为e的原点得到的向量空间，并用ea表示。因为ea是一个向量空间，为了符合我们的符号约定，我们应该使用符号（使用箭头），而不是ea。然而，为了简单起见，我们坚持使用符号ea。

Thus, an affine space →− is a way of defining a vector space structure on a set of points E, without making a commitment to a fixed origin in E. Nevertheless, as soon as we commit to an origin a in E, we can view E as the vector space Ea. However, we urge the reader to think of E as a physical set of points and of →−E as a set of forces acting on E, rather than reducing E to some isomorphic copy of Rn. After all, points are points, and not vectors! For notational simplicity, we will often denote an affine space), or even by E. The vector space →−E is called the vector space associated with E.  
因此，仿射空间→−是在一组点E上定义向量空间结构的一种方法，而不承诺在E中有固定的原点。然而，只要我们承诺在E中有原点A，我们就可以将E视为向量空间EA。然而，我们敦促读者把E看作一个物理点集，把→−E看作作用于E的一组力，而不是把E简化成RN的一些同构副本。毕竟，点是点，而不是向量！为了便于记法，我们通常表示仿射空间），甚至用e表示。向量空间→−e被称为与e关联的向量空间。

 One should be careful about the overloading of the addition symbol +. Addition is well-defined on vectors, as in u + v; the translate a + u of a point a ∈ E by a vector is also well-defined, but addition of points a + b does not make sense. In this respect, the notation b − a for the unique vector u such that b = a + u is somewhat confusing, since it suggests that points can be subtracted (but not added!).  
注意加法符号+的过载。加在向量上定义得很好，如在u+v中；A点∈e的平移a+u也定义得很好，但是加a+b点没有意义。在这方面，唯一向量u的符号b−a使得b=a+u有些混乱，因为它表明可以减去点（但不能相加！）.

Any vector space →−E has an affine space structure specified by choosing , and letting + be addition in the vector space →−E. We will refer to the affine structure on a vector space →−E as the canonical (or natural) affine structure on →−E. In particular, the vector space Rn can be viewed as the affine space , denoted by An. In general, if K is any field, the affine space is denoted by AnK. In order to distinguish between the double role played by members of Rn, points and vectors, we will denote points by row vectors, and vectors by column vectors. Thus, the action of the vector space Rn over the set Rn simply viewed as a set of points is given by  
任何向量空间→−e都有一个通过选择指定的仿射空间结构，并让+在向量空间→−e中相加。我们将向量空间→−e上的仿射结构称为→−e上的规范（或自然）仿射结构。特别是，向量空间rn可以是视为仿射空间，用表示。一般来说，如果k是任何字段，仿射空间用ank表示。为了区分RN成员、点和向量所扮演的双重角色，我们将用行向量表示点，用列向量表示向量。因此，向量空间Rn在仅视为一组点的集合Rn上的作用由下式给出：

.  
.

We will also use the convention that if x = (x1,...,xn) ∈ Rn, then the column vector associated with x is denoted by x (in boldface notation). Abusing the notation slightly, if a ∈ Rn is a point, we also write a ∈ An. The affine space An is called the real affine space of dimension n. In most cases, we will consider n = 1,2,3.  
我们还将使用惯例，如果x=（x1，…，xn）∈rn，那么与x相关的列向量用x表示（粗体符号）。稍微滥用符号，如果a∈rn是一个点，我们也写a∈an。仿射空间an被称为维n的实仿射空间。在大多数情况下，我们将考虑n=1,2,3。

*L*

Figure 23.5: An affine space: the line of equation x + y − 1 = 0.  
图23.5：仿射空间：方程x+y−1=0的直线。

## 23.2 Examples of Affine Spaces 23.2仿射空间示例

Let us now give an example of an affine space that is not given as a vector space (at least, not in an obvious fashion). Consider the subset L of A2 consisting of all points (x,y) satisfying the equation  
现在让我们给出一个仿射空间的例子，它不是作为向量空间给出的（至少，不是以明显的方式）。考虑a2的子集l，由满足方程式的所有点（x，y）组成。

x + y − 1 = 0.  
X+Y−1=0。

The set L is the line of slope −1 passing through the points (1,0) and (0,1) shown in Figure  
设置L是穿过图中所示点（1,0）和（0,1）的坡度线−1。

23.5.  
23.5。

The line L can be made into an official affine space by defining the action +: L×R → L of R on L defined such that for every point (x,1 − x) on L and any u ∈ R,  
L线可以通过定义L上r的作用+：l×r→l而成为正式的仿射空间，定义为L上的每一点（x，1−x）和任何u∈r，

(x,1 − x) + u = (x + u,1 − x − u).  
（X，1−X）+U=（X+U，1−X−U）。

It is immediately verified that this action makes L into an affine space. For example, for any two points a = (a1,1 − a1) and b = (b1,1 − b1) on L, the unique (vector) u ∈ R such that b = a + u is u = b1 − a1. Note that the vector space R is isomorphic to the line of equation x + y = 0 passing through the origin.  
立即证实这个动作使L变成仿射空间。例如，对于L上的任意两点a=（a1,1−a1）和b=（b1,1−b1），唯一（向量）u∈r使得b=a+u是u=b1−a1。注意，向量空间r与通过原点的方程x+y=0的直线同构。

Similarly, consider the subset H of A3 consisting of all points (x,y,z) satisfying the equation  
同样，考虑a3的子集h，由满足方程的所有点（x，y，z）组成。

x + y + z − 1 = 0.  
X+Y+Z−1=0。

The set H is the plane passing through the points (1,0,0), (0,1,0), and (0,0,1). The plane H can be made into an official affine space by defining the action +: H × R2 → H of R2 on  
设置h是穿过点（1,0,0）、（0,1,0）和（0,0,1）的平面。平面h可以通过定义上r2的作用+：h×r2→h而成为正式的仿射空间。

### 23.3. CHASLES’S IDENTITY 23.3。查理斯的身份

(0

,0,

1)

H

Figure 23.6: An affine space: the plane x + y + z − 1 = 0.  
图23.6：仿射空间：平面x+y+z−1=0。

H defined such that for every point (x,y,1 − x − y) on H and any ,  
h的定义是，对于h和任意点（x，y，1−x−y）的每个点，

.  
.

For a slightly wilder example, consider the subset P of A3 consisting of all points (x,y,z) satisfying the equation  
对于一个稍微疯狂的例子，考虑a3的子集p，它由满足方程的所有点（x，y，z）组成。

x2 + y2 − z = 0.  
x2+y2−z=0.

The set P is a paraboloid of revolution, with axis Oz. The surface P can be made into an official affine space by defining the action +: P × R2 → P of R2 on P defined such that for every point (x,y,x2 + y2) on P and any ,  
集合P是一个旋转的抛物面，轴为oz。曲面P可以通过定义p上r2的作用+：p×r2→p而变成一个正式的仿射空间，定义为p和an y上的每一点（x，y，x2+y2）。

.  
.

See Figure 23.7.  
见图23.7。

This should dispel any idea that affine spaces are dull. Affine spaces not already equipped with an obvious vector space structure arise in projective geometry.  
这应该可以消除仿射空间单调的想法。射影几何中出现的仿射空间尚未配备明显的矢量空间结构。

## 23.3 Chasles’s Identity 23.3 Challes的身份

Given any three points a,b,c ∈ E, since c = a + →−ac, b = a + →−ab, and c = b + →−bc, we get  
任意三点a，b，c∈e，因为c=a+→−ac，b=a+→−ab，c=b+→−bc，我们得到

c = b + →−bc = (a + →−ab) + →−bc = a + (→−ab + →−bc)  
C=B+？-BC=（A+？-AB）+？-BC=A+（？-AB+？-BC）

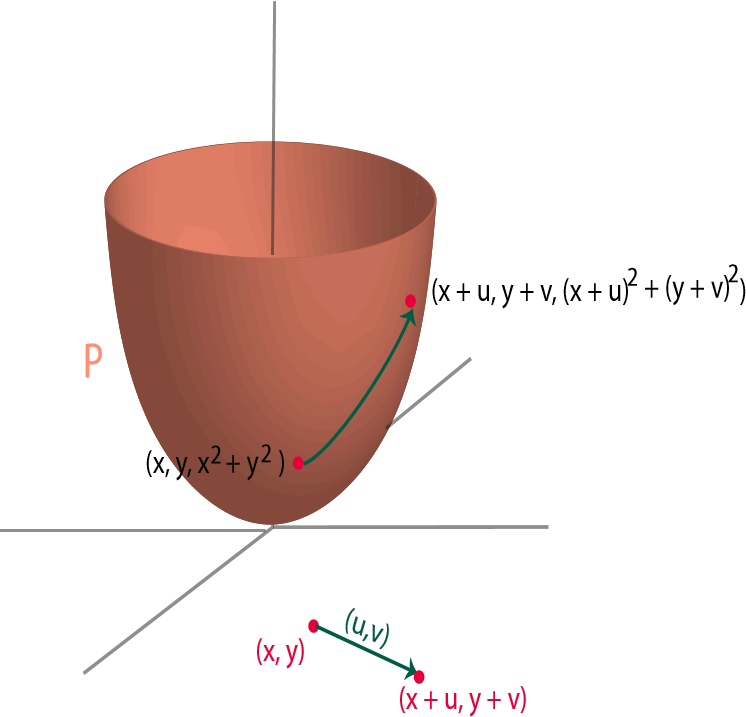


Figure 23.7: The paraboloid of revolution P viewed as a two-dimensional affine space.  
图23.7：旋转p的抛物面被视为二维仿射空间。

by (A2), and thus, by (A3),  
通过（a2），因此，通过（a3）

→−ab + →−bc = →−ac,  
→−AB+→−BC=→−AC，

which is known as Chasles’s identity, and illustrated in Figure 23.8. Since a = a + −aa→ and by (A1) a = a + 0, by (A3) we get  
这被称为Challes的身份，如图23.8所示。由于a=a+−aa→和（a1）a=a+0，由（a3）我们得到

−aa→ = 0.  
−AA→=0。

Thus, letting a = c in Chasles’s identity, we get  
因此，让A=C在Challes的身份中，我们得到

→−ba = −→−ab.  
→−BA=−→−AB.

Given any four points a,b,c,d ∈ E, since by Chasles’s identity  
给定任意四点a，b，c，d∈e，因为由Challes的同一性

→−ab + →−bc = −ad→ + →−dc = →−ac,  
→−AB+→−BC=−AD→→−DC=→−AC，

we have the parallelogram law  
我们有平行四边形定律

→−ab = →−dc iff →−bc = −ad.→  
→−AB=→−DC IFF→−BC=−AD.→

## 23.4 Affine Combinations, Barycenters 23.4仿射组合，重心

A fundamental concept in linear algebra is that of a linear combination. The corresponding concept in affine geometry is that of an affine combination, also called a barycenter. However, there is a problem with the naive approach involving a coordinate system, as we saw in Section 23.1. Since this problem is the reason for introducing affine combinations, at the  
线性代数的一个基本概念是线性组合。仿射几何中相应的概念是仿射组合，也称为重心。但是，我们在23.1节中看到，涉及坐标系的幼稚方法存在问题。因为这个问题是引入仿射组合的原因，在

### 23.4. AFFINE COMBINATIONS, BARYCENTERS 23.4。仿射组合，重心

*E*

−

→

*E*

*a*

*b*

*c*

−

→

*ab*

−

→

*bc*

−

→

*ac*

Figure 23.8: Points and corresponding vectors in affine geometry.  
图23.8：仿射几何中的点和相应的向量。

risk of boring certain readers, we give another example showing what goes wrong if we are not careful in defining linear combinations of points.  
如果我们不小心定义点的线性组合，我们会给出另一个例子，说明出了什么问题。

Consider R2 as an affine space, under its natural coordinate system with origin O = (0,0) and basis vectors and . Given any two points a = (a1,a2) and b = (b1,b2), it is natural to define the affine combination λa + µb as the point of coordinates  
把r2看作是仿射空间，在它的自然坐标系下，原点为o=（0,0），基向量为和。任意两点a=（a1，a2）和b=（b1，b2），将仿射组合λa+μb定义为坐标点是很自然的。

(λa1 + µb1,λa2 + µb2).  
（λa1+μb1，λa2+μb2）。

Thus, when a = (−1,−1) and b = (2,2), the point a + b is the point c = (1,1).  
因此，当a=−1、−1）和b=（2,2）时，点a+b是点c=（1,1）。

Let us now consider the new coordinate system with respect to the origin c = (1,1) (and the same basis vectors). This time, the coordinates of a are (−2,−2), the coordinates of b are (1,1), and the point a+b is the point d of coordinates (−1,−1). However, it is clear that the point d is identical to the origin O = (0,0) of the first coordinate system. This situation is illustrated in Figure 23.9.  
现在让我们考虑新的坐标系关于原点c=（1,1）（和相同的基向量）。这一次，a的坐标是（−2、−2），b的坐标是（1,1），a+b是坐标的点d（−1、−1）。但是，很明显，点D与第一个坐标系的原点O=（0,0）相同。这种情况如图23.9所示。

Thus, a+b corresponds to two different points depending on which coordinate system is used for its computation!  
因此，A+B对应于两个不同的点，这取决于用于计算的坐标系！

This shows that some extra condition is needed in order for affine combinations to make sense. It turns out that if the scalars sum up to 1, the definition is intrinsic, as the following proposition shows.  
这表明为了使仿射组合有意义，需要一些额外的条件。事实证明，如果scalars加起来是1，那么这个定义是内在的，如下所示。

Proposition 23.1. Given an affine space E, let (ai)i∈I be a family of points in E, and let (λi)i∈I be a family of scalars. For any two points a,b ∈ E, the following properties hold:  
提案23.1.给定仿射空间e，设（a i）i∈i为e中的点族，设（λi）i∈i为标量族。对于任意两点a，b∈e，下列性质成立：

1. If Pi∈I λi = 1, then  
   如果pi∈iλi=1，则

.  
.

O = (0,0)

a = (-1,-1)

b = (2,2)

c = a + b = (1,1)

c

a = (-2, -2)

b = (1,1)

d = a + b = (-1,-1)

Figure 23.9: The example from the beginning of Section 23.4.  
图23.9：第23.4节开头的示例。

1. If Pi∈I λi = 0, then X X −→ λiaa−→i = λibai.  
   如果pi∈iλi=0，则x x−→λiaa−→i=λibai。

i∈I i∈I  
I∈I I∈I

Proof. (1) By Chasles’s identity (see Section 23.3), we have  
证据。（1）根据Challes的身份（见第23.3节），我们

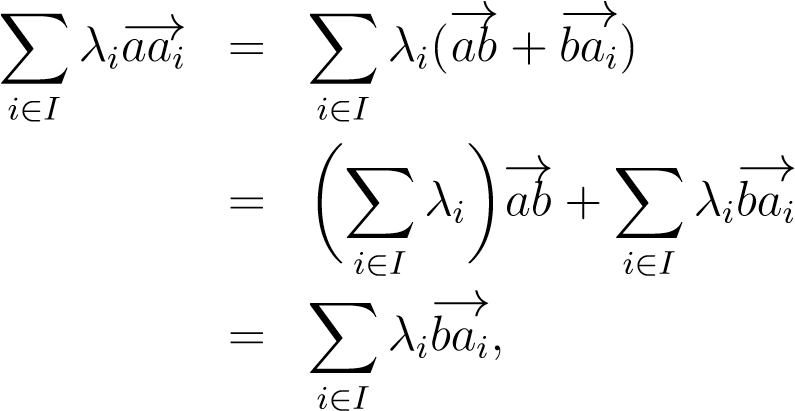
since Pi∈I λi = 1  
因为pi∈iλi=1

= b + Xλiba−→i since b = a + →−ab.  
=b+xλi b a−→i，因为b=a+→−ab。

i∈I  
我爱我

An illustration of this calculation in A2 is provided by Figure 23.10.  
图23.10给出了A2中该计算的图解。

(2) We also have  
（2）我们也有



since Pi∈I λi = 0.   
因为π∈iλi=0。

### 23.4. AFFINE COMBINATIONS, BARYCENTERS 23.4。仿射组合，重心

a

a

a

1

2

3

i

a

a

a

3

a

a

3

a

b

3

a

2

a

a

2

a

b

2

a

a

1

b

a

1

a

1

a

i

a

i

a

a

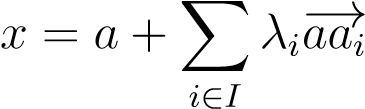
ab

ab

b b  
乙乙

Figure 23.10: Part (1) of Proposition 23.1.  
图23.10：提案23.1的第（1）部分。

Thus, by Proposition 23.1, for any family of points (ai)i∈I in E, for any family (λi)i∈I of scalars such that Pi∈I λi = 1, the point  
因此，通过命题23.1，对于任意点族（ai）i∈i in e，对于任意点族（λi）i∈i的标量，使得pi∈iλi=1，点



is independent of the choice of the origin a ∈ E. This property motivates the following definition.  
独立于原点a∈e的选择，该性质激发了以下定义。

Definition 23.2.P λi = 1, and for anyFor any family of points (a ∈ E, the pointai)i∈I in E, for any family (λi)i∈I of scalars such that i∈I  
定义23.2.pλi=1，对于任意点族（a∈e，点ai）i∈i in e，对于任意点族（λi）i∈i of scalars，使i∈i

a + Xλiaa−→i  
A+Xλi a a−→I

i∈I  
我爱我

(which is independent of a ∈ E, by Proposition 23.1) is called the barycenter (or barycentric combination, or affine combination) of the points ai assigned the weights λi, and it is denoted by  
（独立于a∈e，由23.1命题）被称为分配给权重λi的点ai的重心（或重心组合或仿射组合），并用表示

X  
X

λiai.  
LAI IAI。

i∈I  
我爱我

In dealing with barycenters, it is convenient to introduce the notion of a weighted point, which is just a pair (a,λ), where a ∈ E is a point, and λ ∈ R is a scalar. Then, given a family of weighted points ((ai,λi))i∈I, where Pi∈I λi = 1, we also say that the point Pi∈I λiai is the barycenter of the family of weighted points ((ai,λi))i∈I.  
在处理重心问题时，引入一个加权点的概念是很方便的，它只是一对（a，λ），其中a∈e是一个点，而λ∈r是一个标量。然后，给定一个加权点族（（a i，λi））i∈i，其中pi∈iλi=1，我们也可以说点pi∈iλiai是加权点族的重心（（ai，λi））i∈i。

Note that the barycenter x of the family of weighted points ((ai,λi))i∈I is the unique point such that  
注意，加权点族的重心x（（ai，λi））i∈i是唯一的点，因此

for every a ∈ E,  
对于每一个a∈e，

and setting a = x, the point x is the unique point such that  
设置a=x，点x是唯一的点，这样

Xλixa−→i = 0.  
xλixa－→i=0.

i∈I  
我爱我

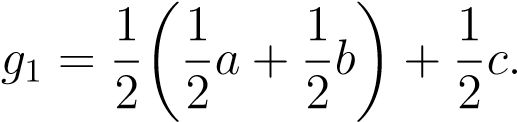
In physical terms, the barycenter is the center of mass of the family of weighted points ((ai,λi))i∈I (where the masses have been normalized, so that Pi∈I λi = 1, and negative masses are allowed).  
在物理术语中，重心是加权点族（（ai，λi））i∈i的质量中心（其中质量已经归一化，使得pi∈iλi=1，并且允许负质量）。

Remarks:  
评论：

1. (Since the barycenter of a family ((λi)i∈I of scalars with finite support (and such thatai,λi))i∈I of weighted points is defined for familiesPi∈I λi = 1), we might as well assume that I is finite. Then, for all m ≥ 2, it is easy to prove that the barycenter of m weighted points can be obtained by repeated computations of barycenters of two weighted points.  
   （由于有有限支持的标量族（（λi）i∈i（并且这样ai，λi））i∈i的加权点是为族定义的，所以我们也可以假设i是有限的。然后，对于所有m≥2，通过对两个加权点的重心的重复计算，很容易证明m加权点的重心是可以得到的。
2. This result still holds, provided that the field K has at least three distinct elements, but the proof is trickier!  
   这个结果仍然成立，前提是K域至少有三个不同的元素，但是证明更难！
3. denote it byWhen Pi∈I λPi = 0i∈I λ, the vectoriai. This observation will be used to define a vector space in whichPi∈I λiaa−→i does not depend on the point a, and we may linear combinations of both points and vectors make sense, regardless of the value of Pi∈I λi.  
   当pi∈iλpi=0i∈iλ时，用向量表示。这个观察将被用来定义一个向量空间，其中pi∈iλi a a−→i不依赖于a点，我们可以使点和向量的线性组合有意义，不管pi∈iλi的值如何。

Figure 23.11 illustrates the geometric construction of the barycenters g1 and g2 of the weighted points a, , and c, , and (a,−1), (b,1), and (c,1).  
图23.11说明了加权点A、C和（A、−1）、（B、1）和（C、1）的重心g1和g2的几何结构。

The point g1 can be constructed geometrically as the middle of the segment joining c to the middle of the segment (a,b), since  
点g1可以几何构造为连接c到段（a，b）中间的段的中间，因为



The point g2 can be constructed geometrically as the point such that the middle of the segment (b,c) is the middle of the segment (a,g2), since  
点g2可以几何构造为点，使得段（b，c）的中间是段（a，g2）的中间，因为

.  
.

Later on, we will see that a polynomial curve can be defined as a set of barycenters of a fixed number of points. For example, let (a,b,c,d) be a sequence of points in A2. Observe that  
稍后，我们将看到一条多项式曲线可以定义为一组具有固定数量点的重心。例如，让（A、B、C、D）是A2中的点序列。注意

(1 − t)3 + 3t(1 − t)2 + 3t2(1 − t) + t3 = 1,  
（1−t）3+3t（1−t）2+3t 2（1−t）+t3=1，

*a*

*b*

*c*

*g*

1

*a*

*b*

*c*

*g*

2

Figure 23.11: Barycenters,  
图23.11：重心，

since the sum on the left-hand side is obtained by expanding (t + (1 − t))3 = 1 using the binomial formula. Thus,  
因为左侧的和是通过使用二项式展开（t+（1-t））3=1得到的。因此，

(1 − t)3 a + 3t(1 − t)2 b + 3t2(1 − t)c + t3 d  
（1−T）3 A+3T（1−T）2 B+3T2（1−T）C+T3 D

is a well-defined affine combination. Then, we can define the curve F : A → A2 such that  
是定义明确的仿射组合。然后，我们可以定义曲线f:a→a2，这样

F(t) = (1 − t)3 a + 3t(1 − t)2 b + 3t2(1 − t)c + t3 d.  
F（t）=（1−t）3 A+3T（1−t）2 B+3T2（1−t）C+T3 D。

Such a curve is called a B´ezier curve, and (a,b,c,d) are called its control points. Note that the curve passes through a and d, but generally not through b and c. It can be sbown that any point F(t) on the curve can be constructed using an algorithm performing affine interpolation steps (the de Casteljau algorithm).  
这样的曲线称为B’ezier曲线，并且（A、B、C、D）称为其控制点。注意曲线通过A和D，但一般不通过B和C。曲线上的任何点F（T）都可以使用执行仿射插值步骤的算法（de casteljau算法）来构造。

## 23.5 Affine Subspaces 23.5仿射子空间

In linear algebra, a (linear) subspace can be characterized as a nonempty subset of a vector space closed under linear combinations. In affine spaces, the notion corresponding to the notion of (linear) subspace is the notion of affine subspace. It is natural to define an affine subspace as a subset of an affine space closed under affine combinations.  
在线性代数中，（线性）子空间可以表示为在线性组合下闭合的向量空间的非空子集。在仿射空间中，与（线性）子空间概念相对应的概念是仿射子空间的概念。将仿射子空间定义为仿射组合下封闭的仿射空间的子集是很自然的。

Definition 23.3. Given an affine space, a subset V of E is an affine subspace (of  
定义23.3.给定仿射空间，e的子集v是仿射子空间（of

) if for every family of weighted pointsP λiai belongs to V . ((ai,λi))i∈I in V such that Pi∈I λi = 1, the barycenter i∈I  
）如果每个加权点族的λi ai都属于v.（（ai，λi））i∈i在v中使得pi∈iλi=1，则重心i∈i

An affine subspace is also called a flat by some authors. According to Definition 23.3, the empty set is trivially an affine subspace, and every intersection of affine subspaces is an affine subspace.  
仿射子空间也被一些作者称为平面。根据23.3的定义，空集通常是仿射子空间，仿射子空间的每个交叉点都是仿射子空间。

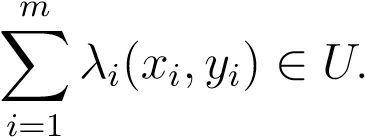
As an example, consider the subset U of R2 defined by  
例如，考虑由

,  
，

i.e., the set of solutions of the equation  
即方程组的解

ax + by = c,  
ax+by=c，

where it is assumed that a = 06 or b = 06 . Given any m points (xi,yi) ∈ U and any m scalars λi such that λ1 + ··· + λm = 1, we claim that  
假设a=06或b=06。给定任何m点（十一，易）u和任何m个标量i i，使之等于，1 + +，+，m＝1，我们声称

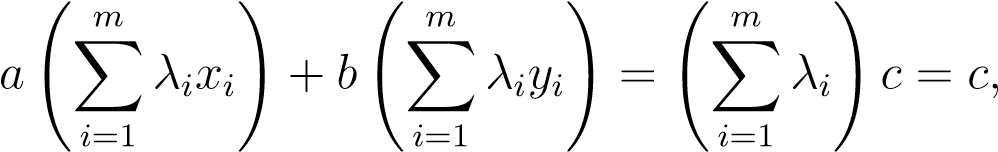


Indeed, (xi,yi) ∈ U means that axi + byi = c,  
实际上，（Xi，Yi）u表示AXI+BYI= C，

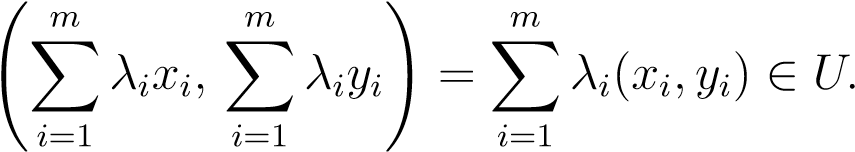
and if we multiply both sides of this equation by λi and add up the resulting m equations, we get  
如果我们用这个方程的两边乘以λi，把得到的m方程加起来，我们得到



and since λ1 + ··· + λm = 1, we get  
既然λ1+·····+λm=1，我们得到



which shows that  
这表明



Thus, U is an affine subspace of A2. In fact, it is just a usual line in A2. It turns out that U is closely related to the subset of R2 defined by  
因此，u是a2的仿射子空间。实际上，它只是A2中的一条普通线。结果表明，u与由

,  
，

i.e., the set of solutions of the homogeneous equation  
即齐次方程的一组解

ax + by = 0  
ax+by=0

*U*

−

→

*U*

Figure 23.12: An affine line U and its direction.  
图23.12：仿射线U及其方向。

obtained by setting the right-hand side of ax + by = c to zero. Indeed, for any m scalars λi, the same calculation as above yields that  
通过将ax+的右侧x=c设置为零获得。事实上，对于任何m标量λi，与上面相同的计算得出：

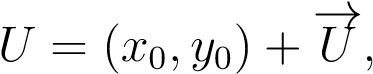
m  
米

X ∈ →−  
X∈→−

λi(xi,yi) U ,  
（I，I）U，

i=1  
i＝1

this time without any restriction on the , since the right-hand side of the equation is null. Thus, →−U is a subspace of R2. In fact, U is one-dimensional, and it is just a usual line in R2. This line can be identified with a line passing through the origin of A2, a line that is parallel to the line U of equation ax + by = c, as illustrated in Figure 23.12. Now, if (x0,y0) is any point in U, we claim that  
这一次对没有任何限制，因为方程的右边是空的。因此，→−u是r2的一个子空间。实际上，u是一维的，它只是r2中的一条普通线。这条线可以用穿过a2原点的线来标识，a2原点与方程ax+by=c的u线平行，如图23.12所示。现在，如果（X0，Y0）是U中的任何一个点，我们声称



where (x0,y0) + →−U = (x0 + u1,y0 + u2) | (u1,u2) ∈ →−U . n o  
其中（X0，Y0）+→−U=（X0+U1，Y0+U2）（U1，U2）∈→−U。氮氧化物

First, (x0,y0) + →−U ⊆ U, since ax0 + by0 = c and au1 + bu2 = 0 for all (u1,u2) ∈ →−U . Second, if (x,y) ∈ U, then ax + by = c, and since we also have ax0 + by0 = c, by subtraction, we get  
首先，（X0，y0）+→−u u，因为ax0+by0=c，au1+bu2=0代表所有（u1，u2）∈→−u。其次，如果（x，y）∈u，那么ax+by=c，由于我们也有ax0+by0=c，通过减法，我们得到

a(x − x0) + b(y − y0) = 0,  
A（X−X0）+B（Y−Y0）=0，

which shows that (→− x − x0,y − , and thus (x,y) ∈ (x0,y0) + →−U . Hence, we also have U ⊆ (x0,y0) + U , and U = (x0,y0) + U .  
这表明（→−X−X0，Y−，因此（X，Y）∈（X0，Y0）+→−U。因此，我们也有u（x0，y0）+u和u=（x0，y0）+u。

The above example shows that the affine line U defined by the equation  
上面的例子表明，方程定义的仿射线u

ax + by = c  
ax+by=c

is obtained by “translating” the parallel line →−U of equation  
通过“平移”方程的平行线→−U得到

ax + by = 0  
ax+by=0

passing through the origin. In fact, given any point (x0,y0) ∈ U,  
穿过原点。实际上，给定任意点（X0，Y0）∈U，

U = (x0,y0) + →−U .  
U=（X0，Y0）+→−U。

More generally, it is easy to prove the following fact. Given any m × n matrix A and any vector b ∈ Rm, the subset U of Rn defined by  
一般来说，很容易证明以下事实。给定任意m×n矩阵a和任意向量b∈rm，由

U = {x ∈ Rn | Ax = b}  
u=x∈rn ax=b

is an affine subspace of An.  
是的仿射子空间。

Actually, observe that Ax = b should really be written as Ax> = b, to be consistent with our convention that points are represented by row vectors. We can also use the boldface notation for column vectors, in which case the equation is written as Ax = b. For the sake of minimizing the amount of notation, we stick to the simpler (yet incorrect) notation Ax = b.  
实际上，注意ax=b应该写成ax>=b，以符合我们的惯例，即点由行向量表示。我们也可以对列向量使用黑体表示法，在这种情况下，方程写为ax=b。为了最小化表示法的数量，我们坚持使用更简单（但不正确）的表示法ax=b。

If we consider the corresponding homogeneous equation Ax = 0, the set  
如果我们考虑相应的齐次方程ax=0，

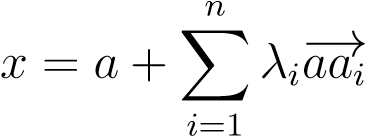
→−U = {x ∈ Rn | Ax = 0}  
→−u=x∈rn ax=0

is a subspace of Rn, and for any x0 ∈ U, we have  
是Rn的子空间，对于任何X0∈U，我们有

U = x0 + →−U .  
U=X0+→−U。

This is a general situation. Affine subspaces can be characterized in terms of subspaces of  
这是一般情况。仿射子空间可以用

→−E. Let V be a nonempty subset of E. For every family (a1,...,an) in V , for any family (λ1,...,λn) of scalars, and for every point a ∈ V , observe that for every x ∈ E,  
→−e.设v为e的非空子集，对于v中的每个族（a1，…，an），对于scalars的任何族（λ1，…，λn），对于每个点a∈v，观察每个x∈e，



is the barycenter of the family of weighted points  
是加权点族的重心

,  
，

since  
自从

.  
.

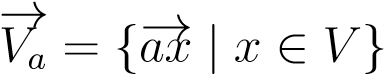
Given any point a ∈ E and any subset →−V of →−E, let a+→−V denote the following subset of E:  
给定任意点a∈e和→−e的任何子集→−v，让a+→−v表示e的以下子集：

a + →−V = na + v | v ∈ →−V o.  
A+→−V=NA+V V∈→−V O.

Figure 23.13: An affine subspace V and its direction →−V .  
图23.13：仿射子空间V及其方向→−V。

Proposition 23.2. Let be an affine space.  
提案23.2.设为仿射空间。

1. A nonempty subset V of E is an affine subspace iff for every point a ∈ V , the set  
   e的一个非空子集v是每个点a∈v的仿射子空间iff，集合



is a subspace of →−E. Consequently, . Furthermore,  
是→−e的子空间。因此，。此外，

is a subspace of E and Va = V for all a ∈ E. Thus,.  
是所有a∈e的e和va=v的子空间，因此。

1. For any subspace →−V of →−E and for any a ∈ E, the set V = a+→−V is an affine subspace.  
   对于→−e的任意子空间→−v，对于任意a∈e，集合v=a+→−v是仿射子空间。

Proof. The proof is straightforward, and is omitted. It is also given in Gallier [71].   
证据。证据很直接，被省略了。加利尔文[71]也给出了这一点。

In particular, when E is the natural affine space associated with a vector space →−E, Proposition 23.2 shows that every affine subspace of E is of the form u + →−U , for a subspace →−U of →−E. The subspaces of →−E are the affine subspaces of E that contain 0.  
特别地，当e是与向量空间→−e相关联的自然仿射空间时，命题23.2表明e的每个仿射子空间的形式为u+→−u，对于→−e的子空间→−u，→−e的子空间是包含0的e的仿射子空间。

The subspace →−V associated with an affine subspace is called the direction of V . It is also clear that the map +: V ×→−V → V induced by +: E × E → E confers to →− an affine structure. Figure 23.13 illustrates the notion of affine subspace.  
与仿射子空间相关联的子空间→−v称为v的方向。还清楚地表明，由+：e×e→e诱导的map+：v×→−v→v赋予→−一个仿射结构。图23.13说明了仿射子空间的概念。

By the dimension of the subspace V , we mean the dimension of →−V .  
通过子空间v的维数，我们是指→−v的维数。

An affine subspace of dimension 1 is called a line, and an affine subspace of dimension 2 is called a plane.  
维1的仿射子空间称为直线，维2的仿射子空间称为平面。

An affine subspace of codimension 1 is called a hyperplane (recall that a subspace F of a vector space E has codimension 1 iff there is some subspace G of dimension 1 such that E = F ⊕ G, the direct sum of F and G, see Strang [165] or Lang [106]).  
余维1的仿射子空间称为超平面（回想一下，向量空间e的子空间f具有余维1，如果存在维度1的某些子空间g，则e=f\_g，f和g的直接和，参见strang[165]或lang[106]）。

We say that two affine subspaces U and V are parallel if their directions are identical. Equivalently, since →−U = →−V , we have U = a + →−U and V →−= b + →−U for any a ∈ U and any b ∈ V , and thus V is obtained from U by the translation ab.  
我们说两个仿射子空间u和v方向相同时是平行的。等价地，由于→−u=→−v，我们有u=a+→−u和v→−=b+→−u表示任何a∈u和任何b∈v，因此v是通过翻译ab从u获得的。

In general, when we talk about n points a1,...,an, we mean the sequence (a1,...,an), and not the set {a1,...,an} (the ai’s need not be distinct). →−  
一般来说，当我们讨论n点a1，…，an时，我们指的是序列（a1，…，an），而不是集合a1，…，an（ai不必是不同的）。δ

By Proposition 23.2, a line is specified by a point a ∈ E and a nonzero vector v ∈ E,  
在23.2命题中，直线由点a∈e和非零向量v∈e指定，

i.e., a line is the set of all points of the form a + λv, for λ ∈ R→−. →−  
也就是说，一条直线是A+λv形式的所有点的集合，对于λ∈R→−。→−

We say that three points a,b,c are collinear if the vectors ab and ac are linearly dependent. If two of the points→− a,b,c are distinct, say a =6 b, then there is a unique λ ∈ R such that →−ac = λab, and we define the ratio .  
如果向量AB和AC是线性相关的，我们就说三个点A，B，C是共线的。如果两个点→−a，b，c是不同的，假设a=6b，那么有一个唯一的λ∈r，这样→−a c=λa b，我们定义了比率。

A plane is specified by a point a ∈ E and two linearly independent vectors u,v ∈ →−E, i.e., a plane is the set of all points of the form a + λu + µv, for λ,µ ∈→−R.→−  
平面由点a∈e和两个线性无关向量u，v∈→−e指定，即平面是形式a+λu+μv的所有点的集合，对于λ，μ∈→−r.→−

We say that four points a,b,c,d are coplanar if the vectors ab, ac, and are linearly dependent. Hyperplanes will be characterized a little later.  
我们说，如果向量AB、AC和线性相关，那么四个点A、B、C、D是共面的。稍后将对超平面进行描述。

Proposition 23.3. Given an affine spaceP λiai (where Pi∈I λi , for any family= 1) is the smallest affine subspace(ai)i∈I of points in  
提案23.3.给定一个仿射空间pλi ai（其中，对于任何族=1，pi∈iλi）是中点的最小仿射子空间（ai）i∈i。

E, the set V of barycenters i∈I containing (ai)i∈I.  
e，重心i∈i的集合v包含（ai）i∈i。

Proof. If (ai)i∈I is empty, then V = ∅, because of the condition Pi∈I λi = 1. If (ai)i∈I is nonempty, then the smallest affine subspace containing (P λiai, and thus, it is enough to show that aVi)is closed under affine combina-i∈I must contain the set V of barycenters i∈I tions, which is immediately verified.   
证据。如果（ai）i∈i为空，则v=∅，因为条件pi∈iλi=1。如果（ai）i∈i是非空的，那么含有（pλiai，因此，足以证明avi）的最小仿射子空间在仿射组合a-i∈i下是封闭的，必须包含质心i∈i的集v，并立即得到验证。

Given a nonempty subset S of E, the smallest affine subspace of E generated by S is often denoted by hSi. For example, a line specified by two distinct points a and b is denoted by ha,bi, or even (a,b), and similarly for planes, etc.  
给定e的非空子集s，s生成的e的最小仿射子空间通常用hsi表示。例如，由两个不同的点a和b指定的一条线用ha、bi或偶数（a、b）表示，同样，对于平面等。

Remarks:  
评论：

1. Since it can be shown that the barycenter of n weighted points can be obtained by repeated computations of barycenters of two weighted points, a nonempty subset V of E is an affine subspace iff for every two points a,b ∈ V , the set V contains all barycentric combinations of a and b. If V contains at least two points, then V is an affine subspace iff for any two distinct points a,b ∈ V , the set V contains the line determined by a and b, that is, the set of all points (1 − λ)a + λb, λ ∈ R.  
   由于可以证明N个加权点的重心可以通过两个加权点的重心的重复计算得到，因此E的一个非空子集V是每两个点A，B∈V的仿射子空间iff，集合V包含A和B的所有重心组合。如果v包含S至少两点，则V是任意两点的仿射子空间iff，a，b∈v，集合V包含由a和b确定的直线，即所有点（1−λ）a+λb，λ∈r的集合。
2. This result still holds if the field K has at least three distinct elements, but the proof is trickier!  
   如果字段k至少有三个不同的元素，这个结果仍然有效，但是证明更难！

## 23.6 Affine Independence and Affine Frames 23.6仿射独立性和仿射框架

Corresponding to the notion of linear independence in vector spaces, we have the notion of affine independence. Given a family (ai)i∈I of points in an affine space E, we will reduce the notion of (affine) independence of these points to the (linear) independence of the families (a−−i→aj)j∈(I−{i}) of vectors obtained by choosing any ai as an origin. First, the following proposition shows that it is sufficient to consider only one of these families.  
对应于向量空间中的线性独立概念，我们有仿射独立的概念。给定仿射空间E中点的族（a i）i∈i，我们将这些点的（仿射）独立性的概念简化为选择任意ai作为原点获得的向量族（a−i→a j）j∈（i−i）的（线性）独立性。首先，下面的命题表明，只考虑其中一个家庭是足够的。

Proposition 23.4. Given an affine space →− , let (ai)i∈I be a family of points in  
提案23.4.给定仿射空间→−，让（a i）i∈i是

E. If the family (a−−i→aj)j∈(I−{i}) is linearly independent for some i ∈ I, then (a−−i→aj)j∈(I−{i}) is linearly independent for every i ∈ I.  
e.如果家族（a−i→a j）j∈（i−i）对某些i∈i是线性独立的，那么（a−i→aj）j∈（i−i）对每个i∈i是线性独立的。

Proof. Assume that the family (a−−i→aj)j∈(I−{i}) is linearly independent for some specific i ∈ I.  
证据。假设家族（a−i→a j）j∈（i−i）对于某些特定的i∈i是线性独立的。

Let k ∈ I with k =6 i, and assume that there are some scalars (λj)j∈(I−{k}) such that  
设k∈i，k=6i，并假设有一些标度（λj）j∈（i−k），这样

X λja−−k→aj = 0.  
xλja−k→aj=0.

j∈(I−{k})  
J∈（I−K）

Since  
自从

a−−k→aj = a−−k→ai + a−−i→aj,  
A−K→AJ=A−K→Ai+A−I→AJ，

we have  
我们有

,  
，

and thus  
因此

Since the family (− { } Pa−−i→aj)j∈(I−{λji}= 0) is linearly independent, we must have, which implies that λj = 0 for all j ∈ (λIj − {= 0k}for all). j ∈  
由于家族（−pa−−i→aj）j∈（i−λji=0）是线性独立的，因此我们必须有，这意味着所有j∈（λij−=0K）的λj=0。

(I i,k ) and j∈(I−{k})  
（i i，k）和j∈（i−k）

We define affine independence as follows.  
我们定义仿射独立性如下。

Definition 23.4. Given an affine space, a family (ai)i∈I of points in E is affinely independent if the family (a−−i→aj)j∈(I−{i}) is linearly independent for some i ∈ I.  
定义23.4.给定一个仿射空间，如果（a−i→a j）j∈（i−i）对某些i∈i线性独立，则e中点的族（ai）i∈i是仿射独立的。

*E*

−

→

*E*

*a*

0

*a*

1

*a*

2

−−

→

*a*

0

*a*

1

−−

→

*a*

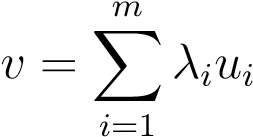
0

*a*

2

Figure 23.14: Affine independence and linear independence  
图23.14：仿射独立和线性独立

Definition 23.4 is reasonable, because by Proposition 23.4, the independence of the family (a−−i→aj)j∈(I−{i}) does not depend on the choice of ai. A crucial property of linearly independent vectors (u1,...,um) is that if a vector v is a linear combination  
定义23.4是合理的，因为根据命题23.4，家庭的独立性（a−i→a j）j∈（i−i）不依赖于人工智能的选择。线性无关向量（u1，…，um）的一个重要特性是，如果向量v是线性组合



of the ui, then the λi are unique. A similar result holds for affinely independent points.  
对于ui，那么λi是唯一的。仿射独立点也有类似的结果。

Proposition 23.5. Given an affine space, let (a0,...,am) be a family of m + 1 points in E. Let x ∈ E, and assume that . Then, the family (λ0,...,λm) such that is unique iff the family is linearly independent.  
提案23.5。给定仿射空间，设（a0，…，a m）为e中m+1点的族，设x∈e，并假定。然后，家族（λ0，…，λm）是唯一的，如果家族是线性独立的。

Proof. The proof is straightforward and is omitted. It is also given in Gallier [71].   
证据。证据很直接，被省略了。加利尔文[71]也给出了这一点。

Proposition 23.5 suggests the notion of affine frame. Affine frames are the affine analogues of bases in vector spaces. Let be a nonempty affine space, and let (a0,...,am) be a family of + 1 points in E. The family (a0,...,am) determines the family of m vectors (. Conversely, given a point a0 in E and a family of m vectors  
命题23.5提出了仿射框架的概念。仿射框架是向量空间中基的仿射类似物。设为非空仿射空间，设（a0，…，a m）为e中+1点的族。族（a0，…，am）决定m向量的族（。相反，给定e中的点a0和m向量族

(u1,...,um) in E, we obtain the family of m+1 points (a0,...,am) in E, where ai = a0 +ui,  
（u1，…，um）在e中，我们得到e中m+1点（a0，…，am）的族，其中ai=a0+ui，

1 ≤ i ≤ m.  
1≤i≤m。

Thus, for any m ≥ 1, it is equivalent to consider a family of→− m + 1 points (a0,...,am) in E, and a pair (a0,(u1,...,um)), where the ui are vectors in E. Figure 23.14 illustrates the notion of affine independence.  
因此，对于任何m≥1，它等价于考虑e中的→−m+1点（a0，…，am）族和一对（a0，（u1，…，um）），其中ui是e中的向量。图23.14说明了仿射独立的概念。

Remark: The above observation also applies to infinite families (ai)i∈I of points in E and families (ui)i∈I−{0} of vectors in →−E, provided that the index set I contains 0.  
注：上述观测也适用于E点的无限族（ai）i∈i和→−e中向量的族（ui）i∈i−0，前提是索引集i包含0。

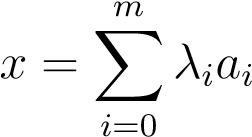
When () is a basis of →−E then, for every x ∈ E, since , there is a unique family (x1,...,xm) of scalars such that  
当（）是→−e的基础时，那么，对于每个x∈e，因为有一个唯一的标量族（x1，…，xm），这样

.  
.

The scalars (x1,...,xm) may be considered as coordinates with respect to )). Since  
scalars（x1，…，xm）可被视为相对于）的坐标。自从

iff ,  
iff ,

x ∈ E can also be expressed uniquely as  
x∈e也可以唯一地表示为



with = 1, and where, and λi = xi for 1 ≤ i ≤ m. The scalars (λ0,...,λm) are also certain kinds of coordinates with respect to (a0,...,am). All this is summarized in the following definition.  
当＝1时，在1和±m＝m的情况下，以及（i，0，…，αm）也是相对于（a0，…，AM）的某种坐标。所有这些在下面的定义中进行了总结。

Definition 23.5. Given an affine space, an affine frame with origin a0 is a family (a0,...,am) of m + 1 points in E such that the list of vectors () is a basis of  
定义23.5.给定仿射空间，原点为a0的仿射帧是e中m+1点的族（a0，…，am），因此矢量列表（）是

→−E. The pair (a0,(a−−0→a1,...,a−−0a→m)) is also called an affine frame with origin a0. Then, every x ∈ E can be expressed as  
→e.这对（a0，（a−0→a1，…，a−0a→m））也被称为原点为a0的仿射帧。那么，每个x∈e可以表示为



for a unique family (x1,...,xm) of scalars, called the coordinates of x w.r.t. the affine frame )). Furthermore, every x ∈ E can be written as  
对于标量的唯一族（x1，…，xm），称为x w.r.t.仿射帧的坐标）。此外，每个x∈e可以写成

x = λ0a0 + ··· + λmam  
x=λ0a0+····+λmam

for some unique family (λ0,...,λm) of scalars such that λ0+···+λm = 1 called the barycentric coordinates of x with respect to the affine frame (a0,...,am). See Figure 23.15.  
对于某些特殊的标度族（λ0，…，λm），如λ0+·····+λm=1，就仿射框架（a0，…，am）而言，称为x的重心坐标。见图23.15。

The coordinates (x1,...,xm) and the barycentric coordinates (λ0,..., λm) are related by the equations and λi = xi, for 1 ≤ i ≤ m. An affine frame is called an affine basis by some authors. A family (ai)i∈I of points in E is affinely dependent if it is not affinely independent. We can also characterize affinely dependent families as follows.  
坐标（x1，…，xm）和重心坐标（α0，…，αm）与方程和Li i＝Xi有关，对于1个i i＝m。仿射框架被一些作者称为仿射基。如果e中点的族（a i）i∈i不是仿射独立的，则它是仿射依赖的。我们还可以将仿射相依族定义为如下。

O

a

0

1)

(1,2,

=

a

1

=

(2,3,

1)

a

2

1)

(-1,3,

=

a

3

=

(1

,3,

2)

x = (-1, 0,2)

O

a

0

a

1

(2,3,

=

1)

a

2

(-1,3,

1)

=

a

3

=

,3,

2)

(1

x = (-1, 0,2)

Figure 23.15: The affine frame (a0,a1,a2,a3) for A3. The coordinates for x = (−1,0,2)  
图23.15:a3的仿射框（a0、a1、a2、a3）。X的坐标=（-1,0,2）

= 1, while the barycentric coordinates for x are λ0 = 3,  
=1，而x的重心坐标为λ0=3，

1 = −  
1=-

Proposition 23.6. Given an affine space , let (ai)i∈I be a family of points in E. jThe family∈ I, i∈ (ai)i∈I is affinely dependent iff there is a familyP λixa−→i = 0 for every x ∈ E. (λi)i∈I such that λj = 06 for some  
提案23.6.给定一个仿射空间，让（a i）i∈i是e中的点族，j该族∈i，i∈（ai）i∈i是仿射相依的，如果存在一个家族，则每个x∈e有一个家族，即λi x a−→i=0。（λi）i∈i，这样对于某些人来说，λj=06

P λi = 0, and i∈I I  
pλi=0，i∈i i

Proof. By Proposition 23.5, the family (ai)i∈I is affinely dependent iff the family of vectors (a−−i→aj)j∈(I−{i}) is linearly dependent for some i ∈ I. For any i ∈ I, the family (a−−i→aj)j∈(I−{i}) is linearly dependent iff there is a family (λj)j∈(I−{i}) such that λj = 06 for some j, and such that  
证据。根据23.5，家族（a i）i∈i是仿射依赖的，如果向量家族（a−i→a j）j∈（i−i）对某些i∈i是线性依赖的，对于任何i∈i，家族（a−i→aj）j∈（i−i）是线性依赖的，如果存在家族（λj）j∈（i−i），那么λj=06 f或者一些J，这样

X λja−−i→aj = 0.  
xλja−i→aj=0.

j∈(I−{i})  
J∈（I−I）

Then, for any x ∈ E, we have  
那么，对于任何x∈e，我们有

,  
，

and letting, we get Pi∈I λixa−→i = 0, with Pi∈I λi = 0 and λj = 06 for some j ∈ I. The converse is obvious by setting x = ai for some i such that λi = 06 , since Pi∈I λi = 0 implies that λj = 06 , for some j =6 i.   
假设，我们得到了π∈iλixa−→i=0，对于一些j∈i，用π∈i=0和λj=06。通过为一些i设置x=ai，使得λi=06明显相反，因为π∈iλi=0意味着对于一些j=6i，用π∈iλi=06。

a2  
A2

a0   
A0

a0a1   
A0A1

Figure 23.16: Examples of affine frames and their convex hulls.  
图23.16：仿射框架及其凸面外壳的示例。

Even though Proposition 23.6 is rather dull, it is one of the key ingredients in the proof of beautiful and deep theorems about convex sets, such as Carath´eodory’s theorem, Radon’s theorem, and Helly’s theorem.  
尽管23.6命题相当单调，但它是证明凸集美丽而深刻定理的关键要素之一，如Carath'Eodory定理、Radon定理和Helly定理。

A family of two points (a,b) in E is affinely independent iff →−ab = 06 , iff a =6 b. If a =6 b, the affine subspace generated by a and b is the set of all points (1−λ)a+λb, which is the unique line passing through a and b. A family of three points (a,b,c) in E is affinely independent iff →−ab and →−ac are linearly independent, which means that a, b, and c are not on the same line (they are not collinear). In this case, the affine subspace generated by (a,b,c) is the set of all points (1 − λ − µ)a + λb + µc, which is the unique plane containing→− →− −→ a, b, and c. A family of four points (a,b,c,d) in E is affinely independent iff ab, ac, and ad are linearly independent, which means that a, b, c, and d are not in the same plane (they are not coplanar). In this case, a, b, c, and d are the vertices of a tetrahedron. Figure 23.16 shows affine frames and their convex hulls for |I| = 0,1,2,3.  
e中的两点族（a，b）是仿射独立的iff→−ab=06，iff a=6b，如果a=6b，a和b生成的仿射子空间是所有点（1−λ）a+λb的集合，是通过a和b的唯一线，e中的三点族（a，b，c）是仿射独立的iff。→−AB和→−AC是线性独立的，这意味着A、B和C不在同一条线上（它们不是共线）。在这种情况下，由（a，b，c）生成的仿射子空间是所有点（1−λ−μ）a+λb+μc的集合，这是包含→−→→−−a，b和c的唯一平面。e中的四个点（a，b，c，d）的族是非仿射独立的iff ab，ac和ad是线性独立的，而mea是线性独立的。n表示a、b、c和d不在同一平面上（它们不是共面的）。在这种情况下，a、b、c和d是四面体的顶点。图23.16显示了i=0,1,2,3的仿射框架及其凸壳。

Given n+1 affinely independent points (a0,...,an) in E, we can consider the set of points λ0a0 +···+λnan, where λ0 +···+λn = 1 and λi ≥ 0 (λi ∈ R). Such affine combinations are called convex combinations. This set is called the convex hull of (a0,...,an) (or n-simplex spanned by (a0,...,an)). When n = 1, we get the segment between a0 and a1, including a0 and a1. When n = 2, we get the interior of the triangle whose vertices are a0,a1,a2, including boundary points (the edges). When n = 3, we get the interior of the tetrahedron whose vertices are a0,a1,a2,a3, including boundary points (faces and edges). The set  
给定e中的n+1仿射独立点（a0，…，an），我们可以考虑一组点λ0a0+·····+λnan，其中λ0+·····+λn=1且λi≥0（λi∈r）。这种仿射组合称为凸组合。这个集合被称为（a0，…，an）（或N-单纯形的（a0，…，an））的凸壳。当n=1时，我们得到a0和a1之间的段，包括a0和a1。当n=2时，我们得到顶点为a0，a1，a2的三角形的内部，包括边界点（边）。当n=3时，我们得到顶点为a0、a1、a2、a3的四面体内部，包括边界点（面和边）。布景

where 0 ≤ λi ≤ 1 (λi ∈ R)}  
其中0≤λi≤1（λi∈r）

is called the parallelotope spanned by (a0,...,an). When E has dimension 2, a parallelotope is also called a parallelogram, and when E has dimension 3, a parallelepiped. Figure 23.17 shows the convex hulls and associated parallelotopes for |I| = 0,1,2,3.  
被称为（a0，…，an）所跨越的Parallelotope。当e的维数为2时，平行切开体也称为平行四边形；当e的维数为3时，平行六面体称为平行四边形。图23.17显示了i=0,1,2,3的凸壳和相关的平行耳。

a

0

a

0

a

0

a

1

a

1

a

2

a

0

a

0

a

1

a

1

a

2

a

3

a

3

Figure 23.17: Examples of affine frames, convex hulls, and their associated parallelotopes.  
图23.17：仿射框架、凸面外壳及其相关的平行耳。

More generally, we say that a subset V of E is convex if for any two points a,b ∈ V , we have c ∈ V for every point c = (1 − λ)a + λb, with 0 ≤ λ ≤ 1 (λ ∈ R).  
一般来说，我们说e的一个子集v是凸的，如果对于任意两点a，b∈v，对于每一点c=（1−λ）a+λb，我们都有c∈v，其中0≤λ≤1（λ∈r）。

 Points are not vectors! The following example illustrates why treating points as vectors may cause problems. Let a,b,c be three affinely independent points in A3. Any point x in the plane (a,b,c) can be expressed as  
点不是向量！以下示例说明了将点作为向量处理可能会导致问题的原因。让a，b，c是a3中的三个仿射独立点。平面（a，b，c）中的任何点x都可以表示为

x = λ0a + λ1b + λ2c,  
x=λ0a+λ1b+λ2c，

where λ0 + λ1 + λ2 = 1. How can we compute λ0,λ1,λ2? Letting a = (a1,a2,a3), b = (b1,b2,b3), c = (c1,c2,c3), and x = (x1,x2,x3) be the coordinates of a,b,c,x in the standard frame of A3, it is tempting to solve the system of equations  
式中，λ0+λ1+λ2=1。我们如何计算λ0，λ1，λ2？假设a=（a1，a2，a3），b=（b1，b2，b3），c=（c1，c2，c3），x=（x1，x2，x3）是a3标准框架中a，b，c，x的坐标，很容易解出方程组。

.  
.

However, there is a problem when the origin of the coordinate system belongs to the plane  
然而，当坐标系的原点属于平面时，存在一个问题。

(a,b,c), since in this case, the matrix is not invertible! What we should really be doing is to solve the system  
（a，b，c），因为在这种情况下，矩阵是不可逆的！我们真正应该做的是解决这个系统

λ0Oa−→ + λ1Ob−→ + λ2Oc−→ = Ox,−→  
λ0oa−→+λ1ob−→+λ2oc−→=Ox，−→

where O is any point not in the plane (a,b,c). An alternative is to use certain well-chosen cross products.  
其中o是平面以外的任何点（a，b，c）。另一种选择是使用某些精选的交叉产品。

It can be shown that barycentric coordinates correspond to various ratios of areas and volumes; see the problems.  
可以看出，重心坐标对应于面积和体积的不同比例；见问题。

## 23.7 Affine Maps 23.7仿射图

Corresponding to linear maps we have the notion of an affine map. An affine map is defined as a map preserving affine combinations.  
对应于线性映射，我们有仿射映射的概念。仿射映射定义为保留仿射组合的映射。

Definition 23.6. Given two affine spaces and , a function f : E → E0 is an affine map iff for every family ((ai,λi))i∈I of weighted points in E such that Pi∈I λi = 1, we have  
定义23.6.在给定两个仿射空间的情况下，函数f:e→e0是每个族（（a i，λi））的仿射映射iff，e中加权点的i∈i，使得pi∈iλi=1

.  
.

In other words, f preserves barycenters.  
换句话说，F保存重心。

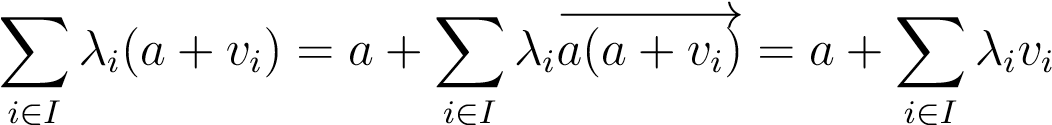
Affine maps can be obtained from linear maps as follows. For simplicity of notation, the same symbol + is used for both affine spaces (instead of using both + and +0).  
仿射映射可以从线性映射中获得，如下所示。为了简化表示法，相同的符号+用于两个仿射空格（而不是同时使用+和+0）。

Proposition 23.7.−→ →Given any point a ∈ E, any point b ∈ E0, and any linear map h: →−E → E0, the map f : E E0 defined such that  
命题23.7。−→→给定任意点a∈e，任意点b∈e0，以及任意线性映射h：→−e→e0，映射f:e e0定义如下：

f(a + v) = b + h(v)  
F（A+V）=B+H（V）

is an affine map.  
是仿射映射。

Proof. Indeed, for any family (λi)i∈I of scalars with Pi∈I λi = 1 and any family (vi)i∈I, since  
证据。事实上，对于任何具有pi∈iλi=1的标量（λi）i∈i和任何族（vi）i∈i，因为



and  
和

,  
，

we have  
我们有

as claimed.   
如要求。

Note that the condition Pi∈I λi = 1 was implicitly used (in a hidden call to Proposition  
注意，条件pi∈iλi=1是隐式使用的（在对命题的隐藏调用中）

23.1) in deriving that  
23.1）在推导

X X  
十倍

λi(a + vi) = a + λivi  
λi（a+vi）=a+λivi

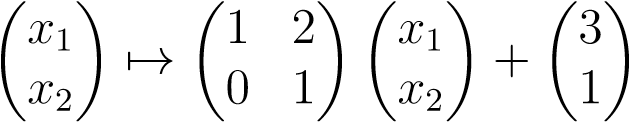
i∈I i∈I  
I∈I I∈I

and  
和

X X λi(b + h(vi)) = b + λih(vi).  
x xλi（b+h（vi））=b+λih（vi）。

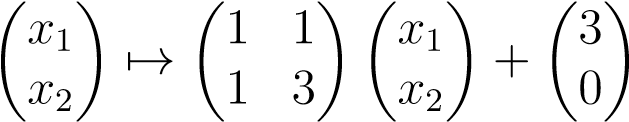
i∈I i∈I  
I∈I I∈I

As a more concrete example, the map  
作为一个更具体的例子，地图



defines an affine map in A2. It is a “shear” followed by a translation. The effect of this shear on the square (a,b,c,d) is shown in Figure 23.18. The image of the square (a,b,c,d) is the parallelogram (a0,b0,c0,d0).  
在A2中定义仿射映射。它是一个“剪切”，然后是一个翻译。这种剪切对正方形（a，b，c，d）的影响如图23.18所示。正方形（a，b，c，d）的图像是平行四边形（a0，b0，c0，d0）。

Let us consider one more example. The map  
让我们再考虑一个例子。地图



is an affine map. Since we can write  
是仿射映射。因为我们可以写

,  
，

d = (5,2) c = (6,2)  
D=（5,2）C=（6,2）

d = (0,1) c = (1,1)   
D=（0,1）C=（1,1）

a = (3,1) b = (4,1)  
A=（3,1）B=（4,1）

a = (0,0) b = (1,0)  
A=（0,0）B=（1,0）

Figure 23.18: The effect of a shear.  
图23.18：剪切效应。

this affine map is the composition of a shear, followed by a rotation of angle π/4, followed by a magnification of ratio √2, followed by a translation. The effect of this map on the square (a,b,c,d) is shown in Figure 23.19. The image of the square (a,b,c,d) is the parallelogram  
这个仿射图是剪切力的组成，然后旋转角度π/4，然后放大比例√2，然后平移。图23.19显示了这张地图对广场（A、B、C、D）的影响。正方形（a，b，c，d）的图像是平行四边形。

(a0,b0,c0,d0).  
（a0、b0、c0、d0）。

* + 1. = (5,4)  
       =（5,4）
    2. = (0,1) c = (1,1)   
       =（0,1）c=（1,1）

a = (0,0) b = (1,0) a = (3,0)  
A=（0,0）B=（1,0）A=（3,0）

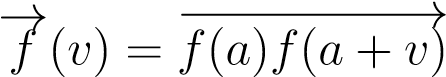
Figure 23.19: The effect of an affine map.  
图23.19：仿射映射的效果。

The following proposition shows the converse of what we just showed. Every affine map is determined by the image of any point and a linear map.  
下面的命题显示了我们刚才所展示的相反的情况。每个仿射映射都由任意点的图像和线性映射决定。

Proposition 23.8. Given an affine map f : E → E0, there is a unique linear map →−f : →−E → such that f(a + v) = f(a) + →−f (v),  
提案23.8。给定仿射映射f:e→e0，有一个唯一的线性映射→−f：→−e→这样f（a+v）=f（a）+→−f（v），

for every a ∈ E and every v ∈ →−E.  
对于每一个a∈e和每一个v∈→−e。

Proof. Let a ∈ E be any point in E. We claim that the map defined such that  
证据。假设a∈e是e中的任意点，我们声称地图定义如下：



for every v ∈ →−E is a linear map →−f : →−E → E−→0. Indeed, we can write  
对于每一个v∈→−e是一个线性映射→−f：→−e→e−→0。事实上，我们可以写

a + λv = λ(a + v) + (1 − λ)a,  
a+λv=λ（a+v）+（1-λ）a，

since, and also  
因为，还有

a + u + v = (a + u) + (a + v) − a,  
A+U+V=（A+U）+（A+V）−A，

since. Since f preserves barycenters, we get  
从那以后。因为F保留了重心，我们得到

f(a + λv) = λf(a + v) + (1 − λ)f(a).  
f（a+λv）=λf（a+v）+（1-λ）f（a）。

If we recall that x = Pi∈I λiai is the barycenter of a family ((ai,λi))i∈I of weighted points (with Pi∈I λi = 1) iff for every b ∈ E,  
如果我们记得x=pi∈iλi ai是一个家族的重心（（ai，λi））i∈i的加权点（pi∈iλi=1）iff对于每个b∈e，

we get  
我们得到

,  
，

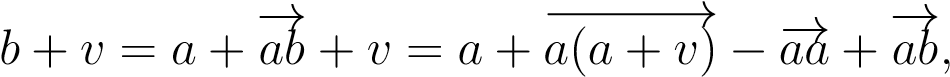
showing that). We also have  
显示出来）。我们也有

f(a + u + v) = f(a + u) + f(a + v) − f(a),  
F（A+U+V）=F（A+U）+F（A+V）−F（A）、

from which we get  
我们从中得到

,  
，

showing that f is a linear map. For any other point b ∈ E, since  
表示f是一个线性映射。对于任何其他点b∈e，因为



b + v = (a + v) − a + b, and since f preserves barycenters, we get  
B+V=（A+V）−A+B，由于F保留了重心，我们得到

f(b + v) = f(a + v) − f(a) + f(b),  
F（B+V）=F（A+V）−F（A）+F（B）

which implies that  
这意味着

,  
，

Thus, f(b)f(b + v) = f(a)f(a + v), which shows that the definition of does not depend on the choice of a ∈ E. The fact that f is unique is obvious: We must have →−f (v) = The unique linear map →−f : →−E → −E→0 given by Proposition 23.8 is called the linear map associated with the affine map f.  
因此，f（b）f（b+v）=f（a）f（a+v），这表明f的定义不依赖于a∈e的选择，f是唯一的这一事实是显而易见的：我们必须有→−f（v）=唯一的线性映射→−f：→−e→−e→0，由23.8给出的，称为与仿射映射。

Note that the condition f(a + v) = f(a) + →−f (v),  
注意，条件f（a+v）=f（a）+→−f（v），

for every a ∈ E and every v ∈ →−E, can be stated equivalently as  
对于每一个a∈e和每一个v∈→−e，可以等价地表示为

f(x) = f(a) + →−f (−ax→), or ,  
f（x）=f（a）+→−f（−ax→），或，

for all a,x ∈ E. Proposition 23.8 shows that for any affine map→−f : →−E → −E→0, such thatf : E → E0, there are points a ∈ E, b ∈ E0, and a unique linear map  
对于所有a，x∈e，命题23.8表明，对于任意仿射映射→−f：→−e→−e→0，这样f:e→e0，有点a∈e，b∈e0，和一个唯一的线性映射

f(a + v) = b + →−f (v),  
F（A+V）=B+→−F（V）

for all v ∈ →−E (just let b = f(a), for any ). Affine maps for which →−f is the identity map are called translations. Indeed, if f = id,  
对于所有v∈→−e（只要让b=f（a），对于任何）。其中→−f是身份图的仿射映射称为翻译。实际上，如果f=id，

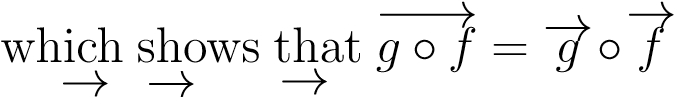
and so   
如此

which shows that f is the translation induced by the vector ) (which does not depend on a).  
这表明f是矢量引起的翻译（不依赖a）。

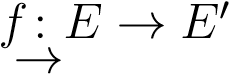
Since an affine map preserves barycenters, and since an affine subspace V is closed under barycentric combinations, the image f(V ) of V is an affine subspace in E0. So, for example, the image of a line is a point or a line, and the image of a plane is either a point, a line, or a plane.  
由于仿射映射保留重心，并且仿射子空间v在重心组合下闭合，因此v的图像f（v）是e0中的仿射子空间。例如，直线的图像是一个点或一条线，平面的图像是一个点、一条线或一个平面。

It is easily verified that the composition of two affine maps is an affine map. Also, given affine maps f : E → E0 and g: E0 → E00, we have  
很容易证明两个仿射映射的组合是一个仿射映射。另外，给定仿射映射f:e→e0和g:e0→e0，我们得到

,  
，



iffAn affine mapf : E → E0 is injective, and thatis constant iff. It is easy to show that an affine mapf :→−fE:→→−E E→0 is surjective iff−E→0 is the null (constant) linear map equal→−f : f→−E: E→→−E→0Eis surjective.0 is injective  
iffan仿射mapf:e→e0是内射的，这是常数iff。可以很容易地看出，仿射映射f：→−f e：→−e→0是可射的iff−e→0是零（常数）线性映射等于→−f:f→−e:e→−e→0是可射的。0是可射的。



to 0 for all v ∈ E.  
对于所有v∈e为0。

If E is an affine space of dimension m and (a0,a1,...,am) is an affine frame for E, then for any other affine space F and for any sequence (b0,b1,...,bm) of m+1 points in F, there is a unique affine map f : E → F such that f(ai) = bi, for 0 ≤ i ≤ m. Indeed, f must be such that  
如果e是维数m的仿射空间，（a0，a1，…，am）是e的仿射框架，那么对于任何其他仿射空间f和f中m+1点的任何序列（b0，b1，…，bm），都有一个唯一的仿射映射f:e→f，因此f（a i）=bi，对于0≤i≤m。实际上，f必须是这样的：

f(λ0a0 + ··· + λmam) = λ0b0 + ··· + λmbm,  
f（λ0a0+····+λmam）=λ0b0+····+λmbm，

where λ0+···+λm = 1, and this defines a unique affine map on all of E, since (a0,a1,...,am) is an affine frame for E.  
其中，λ0+····+λm=1，这定义了所有e上的唯一仿射映射，因为（a0，a1，…，am）是e的仿射帧。

Using affine frames, affine maps can be represented in terms of matrices. We explain how an affine map f : E → E is represented with respect to a frame (a0,...,an) in E, the more general case where an affine map f : E → F is represented with respect to two affine frames (a0,...,an) in E and (b0,...,bm) in F being analogous. Since  
使用仿射框架，仿射映射可以用矩阵表示。我们解释了仿射映射f:e→e是如何相对于e中的帧（a0，…，an）表示的，更一般的情况是仿射映射f:e→f相对于e中的两个仿射帧（a0，…，an）表示，而f中的（b0，…，bm）是类似的。自从

f(a0 + x) = f(a0) + →−f (x)  
F（a0+x）=F（a0）+→−F（x）

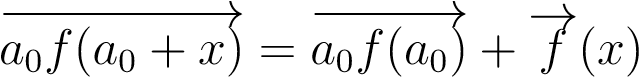
for all x ∈ →−E, we have  
对于所有x∈→−e，我们有

.  
.

Since x, a−−−−0f(a→0), and a−−−−−−−−0f(a0 +→x), can be expressed as  
因为x，a−−−0f（a→0）和a−−−−−0f（a0+→x）可以表示为

|  |  |  |
| --- | --- | --- |
| x 网络错误  a−−−−0f(a→0) 网络错误  a−−−−−−−−0f(a0 +→x) 网络错误 | = 网络错误  = 网络错误  = 网络错误 | x1a−−0→a1 + ··· + xna−−0→an, b1a−−0→a1 + ··· + bna−−0→an, y1a−−0→a1 + ··· + yna−−0→an, 网络错误 |

if A = (aij) is the n×n matrix of the linear map →−f over the basis (, y, and b denote the column vectors of components (x1,...,xn), (y1,...,yn), and (b1,...,bn),  
如果a=（aij）是线性映射的n×n矩阵→−f的基（，y和b表示分量（x1，…，xn），（y1，…，yn）和（b1，…，bn）的列向量，



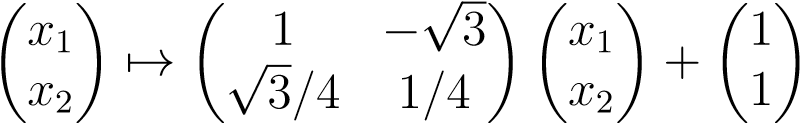
is equivalent to  
等于

y = Ax + b.  
y=ax+b。

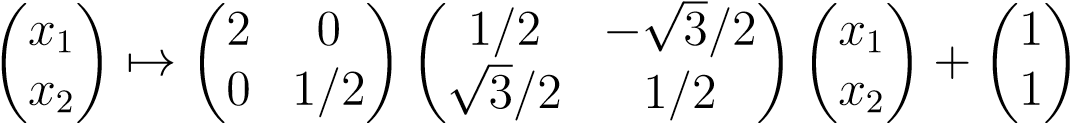
Note that b = 06 unless f(a0) = a0. Thus, f is generally not a linear transformation, unless it has a fixed point, i.e., there is a point a0 such that f(a0) = a0. The vector b is the “translation part” of the affine map. Affine maps do not always have a fixed point. Obviously, nonnull translations have no fixed point. A less trivial example is given by the affine map  
注意b=06，除非f（a0）=a0。因此，f一般不是线性变换，除非它有一个固定点，即有一个点a0，这样f（a0）=a0。向量B是仿射映射的“翻译部分”。仿射映射并不总是有固定点。显然，非空翻译没有固定点。仿射映射给出了一个不那么简单的例子。

.  
.

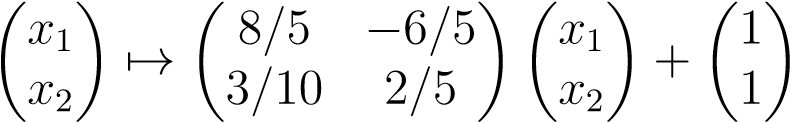
This map is a reflection about the x-axis followed by a translation along the x-axis. The affine map  
这张地图是关于x轴的反射，然后是沿x轴的平移。仿射图



can also be written as  
也可以写为



which shows that it is the composition of a rotation of angle π/3, followed by a stretch (by a factor of 2 along the x-axis, and by a factor of along the y-axis), followed by a translation. It is easy to show that this affine map has a unique fixed point. On the other hand, the affine map  
这表明它是角π/3旋转的组成，然后是拉伸（沿x轴乘以系数2，沿y轴乘以系数），然后是平移。很容易看出这个仿射映射有一个唯一的不动点。另一方面，仿射图

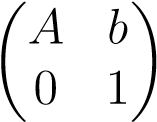


has no fixed point, even though  
没有固定点，即使

,  
，

and the second matrix is a rotation of angle θ such that cos and sin.  
第二个矩阵是角θ的旋转，cos和sin。

There is a useful trick to convert the equation y = Ax + b into what looks like a linear equation. The trick is to consider an (n + 1) × (n + 1) matrix. We add 1 as the (n + 1)th component to the vectors x, y, and b, and form the (n + 1) × (n + 1) matrix  
将方程y=ax+b转换成线性方程有一个很有用的技巧。技巧是考虑（n+1）×（n+1）矩阵。我们将1作为（n+1）th分量加到向量x、y和b上，形成（n+1）×n+1矩阵。



so that y = Ax + b is equivalent to  
所以y=ax+b等于

.  
.

This trick is very useful in kinematics and dynamics, where A is a rotation matrix. Such affine maps are called rigid motions.  
这个技巧在运动学和动力学中非常有用，其中a是旋转矩阵。这种仿射映射称为刚性运动。

If f : E → E0 is a bijective affine map, given any three collinear points a,b,c in E, with a =6 b, where, say, c = (1 − λ)a + λb, since f preserves barycenters, we have f(c) = (1−λ)f(a)+λf(b), which shows that f(a),f(b),f(c) are collinear in E0. There is a converse to this property, which is simpler to state when the ground field is K = R. The converse states that given any bijective function f : E → E0 between two real affine spaces of the same dimension n ≥ 2, if f maps any three collinear points to collinear points, then f is affine. The proof is rather long (see Berger [11] or Samuel [138]).  
如果f:e→e0是双射仿射映射，给定任意三个共线点a，b，c在e中，a=6b，其中，比如，c=（1−λ）a+λb，因为f保留了重心，所以我们得到f（c）=（1−λ）f（a）+λf（b），这表明f（a）、f（b）、f（c）在e0中是共线的。有一个与这个性质相反的性质，当地面场为k=r时，它的状态更简单。在两个相同维n≥2的实仿射空间之间给出了任意双射函数f:e→e0，如果f将任意三个共线点映射到共线点，那么f是仿射的。证据相当长（见Berger[11]或Samuel[138]）。

Given three collinear points a,b,c, where a =6 c, we have b = (1 − β)a + βc for some unique β, and we define the ratio of the sequence a,b, c, as →−  
给定三个共线点a，b，c，其中a=6c，我们得到一些独特β的b=（1−β）a+βc，我们将序列a，b，c的比值定义为→−

ratio(,  
比率（，

provided that β = 16 , i.e., b =6 c. When b = c, we agree that ratio(a,b,c) = ∞. We warn our readers that other authors define the ratio of a,b,c as −ratio( −→. Since affine maps preserve barycenters, it is clear that affine maps preserve the ratio of three points.  
假设β=16，即b=6 c，当b=c时，我们同意比值（a，b，c）=∞。我们警告读者，其他作者将A、B、C的比率定义为−比率（−→）。由于仿射映射保留重心，很明显仿射映射保留了三个点的比例。

## 23.8 Affine Groups 23.8仿射群

We now take a quick look at the bijective affine maps. Given an affine space E, the set of affine bijections f : E → E is clearly a group, called the affine group of , and denoted by GA(E). Recall that the group of bijective linear maps of the vector space E is denoted by GL(→−E). Then, the map f 7→ →−f defines a group homomorphism L: GA(E) → GL(→−E). The kernel of this map is the set of translations on E.  
现在我们快速地看一下双目标仿射图。给定仿射空间e，f:e→e的仿射双射集合显然是一个群，称为仿射群，用ga（e）表示。回想一下，向量空间e的双射线性映射组用gl（→−e）表示。然后，图F7→→−F定义了一个组同态l:ga（e）→gl（→−e）。这个地图的核心是E上的一组翻译。

The subset of all linear maps of the form λid→−E , where λ ∈ R − {0}, is a subgroup of GL(→−E), and is denoted by R∗id→−E (where λid→−E (u) = λu, and R∗ = R − {0}). The subgroup DIL(E) = L−1(R∗id→−E ) of GA(E) is particularly interesting. It turns out that it is the disjoint union of the translations and of the dilatations of ratio λ = 16 . The elements of DIL(E) are called affine dilatations.  
形式为λid→−e的所有线性映射的子集，其中，λ∈r−0，是gl（→−e）的一个子组，并由r id→−e表示（其中，λid→−e（u）=λu，r=r−0）。GA（e）的子组dil（e）=l−1（r id→−e）特别有趣。结果表明，它是平移与扩张之比λ=16的不相交的结合。dil（e）的元素称为仿射扩张。

Given any point a ∈ E, and any scalar λ ∈ R, a dilatation or central dilatation (or homothety) of center a and ratio λ is a map Ha,λ defined such that  
给定任意点a∈e和任意标量λ∈r，中心a和比率λ的扩张或中心扩张（或同构）是映射ha，λ定义如下：

Ha,λ(x) = a + λax,−→  
ha，λ（x）=a+λax，−→

for every x ∈ E.  
每x∈e。

Remark: The terminology does not seem to be universally agreed upon. The terms affine dilatation and central dilatation are used by Pedoe [132]. Snapper and Troyer use the term dilation for an affine dilatation and magnification for a central dilatation [157]. Samuel uses homothety for a central dilatation, a direct translation of the French “homoth´etie” [138]. Since dilation is shorter than dilatation and somewhat easier to pronounce, perhaps we should use that!  
备注：术语似乎没有得到普遍认同。pedoe使用了“仿射扩张”和“中心扩张”这两个术语[132]。Snapper和Troyer使用术语“扩张”表示仿射扩张，而“放大”表示中心扩张[157]。塞缪尔用谐音作为中心扩张词，直接翻译了法语“homoth'etie”[138]。因为扩张比扩张短，而且发音更容易，也许我们应该使用它！

Observe that Ha,λ(a) = a, and when λ = 06 and x =6 a, Ha,λ(x) is on the line defined by a and x, and is obtained by “scaling” ax−→ by λ.  
观察Ha，λ（a）=a，当λ=06和x=6 a时，Ha，λ（x）在a和x定义的线上，并通过“缩放”a x－→由λ获得。

Figure 23.20 shows the effect of a central dilatation of center d. The triangle (a,b,c) is magnified to the triangle (a0,b0,c0). Note how every line is mapped to a parallel line.  
图23.20显示了中心D的中心扩张效应。三角形（a，b，c）放大为三角形（a0，b0，c0）。注意每条线是如何映射到一条平行线的。

When λ = 1, Ha,1 is the identity. Note that. When λ = 06 , it is clear that  
当λ=1时，ha，1为同一性。请注意。当λ=06时，很明显

Ha,λ is an affine bijection. It is immediately verified that  
ha，λ是仿射双射。立即证实

Ha,λ ◦ Ha,µ = Ha,λµ.  
ha，λha，μ=ha，λμ。

We have the following useful result.  
我们得到了以下有用的结果。

### 23.8. AFFINE GROUPS 23.8。仿射群

*d*

*a*

*b*

*c*

*a*

*b*

*c*

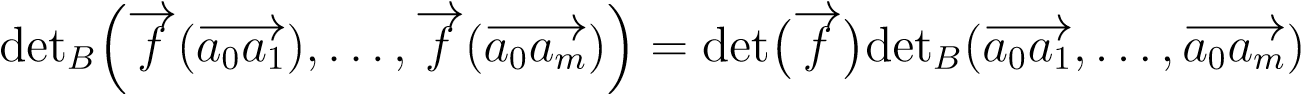
Figure 23.20: The effect of a central dilatation Hd,λ(x).  
图23.20：中心扩张hd的影响，λ（x）。

Proposition 23.9. Given any affine space E, for any affine bijection f ∈ GA(E), if →−f = λid→−E , for some λ ∈ R∗ with λ = 16 , then there is a unique point c ∈ E such that f = Hc,λ.  
提案23.9.给定任意仿射空间e，对于任意仿射双射f∈ga（e），如果→−f=λid→−e，对于某些λ∈r且λ=16，则存在一个唯一点c∈e，使得f=hc，λ。

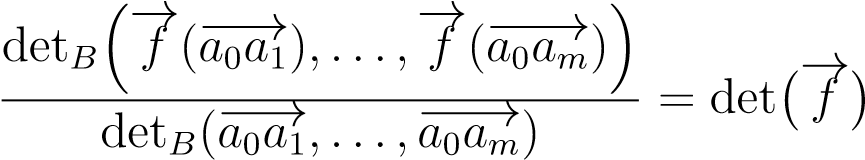
Proof. The proof is straightforward, and is omitted. It is also given in Gallier [71].   
证据。证据很直接，被省略了。加利尔文[71]也给出了这一点。

Clearly, if →−f = id→−E , the affine map f is a translation. Thus, the group of affine dilatations DIL(E) is the disjoint union of the translations and of the dilatations of ratio λ = 06 ,1. Affine dilatations can be given a purely geometric characterization.  
显然，如果→−f=id→−e，仿射映射f就是一个翻译。因此，仿射扩张群dil（e）是平移与比率λ=06，1扩张的不相交的结合。仿射扩张可以给出一个纯粹的几何特征。

Another point worth mentioning is that affine bijections preserve the ratio of volumes of parallelotopes. Indeed, given any basis B = (u1,...,um) of the vector space →−E associated with the affine space E, given any m + 1 affinely independent points (a0,...,am), we can compute the determinant det) w.r.t. the basis B. For any bijective affine map f : E → E, since  
另一点值得一提的是仿射双射保留了平行耳的体积比。实际上，如果向量空间的任何基b=（u1，…，um）→e与仿射空间e相关联，给定任何m+1仿射独立点（a0，…，am），我们可以计算行列式det）w.r.t.对于任何双射仿射映射f:e→e，因为



and the determinant of a linear map is intrinsic (i.e., depends only on →−f , and not on the particular basis B), we conclude that the ratio  
线性映射的行列式是内在的（即只取决于→−f，而不是特定的基础b），我们得出如下结论：



is independent of the basis) is the volume of the parallelotope spanned by (a0,...,am), where the parallelotope spanned by any point a and the vectors  
不依赖于基）是由（a0，…，am）所跨越的平行头的体积，其中平行头由任意点a和向量所跨越。

A

B

a

1

a

2

a

3

b

1

b

2

b

3

H

1

H

2

H

3

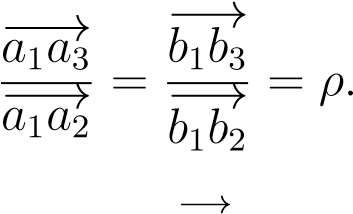
Figure 23.21: The theorem of Thales.  
图23.21：泰雷兹定理。

(u1,...,um) has unit volume (see Berger [11], Section 9.12), we see that affine bijections preserve the ratio of volumes of parallelotopes. In fact, this ratio is independent of the choice of the parallelotopes of unit volume. In particular, the affine bijections f ∈ GA(E) such that det = 1 preserve volumes. These affine maps form a subgroup SA(E) of GA(E) called the special affine group of E. We now take a glimpse at affine geometry.  
（u1，…，um）有单位体积（见Berger[11]，第9.12节），我们发现仿射双射保留了平行耳的体积比。事实上，这一比例与单位体积的平行叶的选择无关。尤其是仿射双射f∈ga（e），使得det=1保留体积。这些仿射映射形成了GA（e）的一个子群sa（e），称为e的特殊仿射群。现在我们来看看仿射几何。

## 23.9 Affine Geometry: A Glimpse 23.9仿射几何：一瞥

In this section we state and prove three fundamental results of affine geometry. Roughly speaking, affine geometry is the study of properties invariant under affine bijections. We now prove one of the oldest and most basic results of affine geometry, the theorem of Thales.  
在这一节中，我们陈述并证明了仿射几何的三个基本结果。一般来说，仿射几何是研究仿射双射下的不变量性质。我们现在证明了仿射几何最古老和最基本的结果之一，泰雷兹定理。

Proposition 23.10. Given any affine space E, if H1,H2,H3 are any three distinct parallel hyperplanes, and A and B are any two lines not parallel to Hi, letting ai = Hi ∩ A and bi = Hi ∩ B, then the following ratios are equal:  
提案23.10。给定任意仿射空间e，如果h1、h2、h3是任意三个不同的平行超平面，而a和b是任意两条不平行于hi的线，让ai=hi a和bi=hi b，则下列比值相等：



Conversely, for any point d on the line, then d = a3.  
相反，对于线上的任何点d，则d=a3。

Proof. Figure 23.21 illustrates the theorem of Thales. We sketch a proof, leaving the details as an exercise. Since H1,H2, H3 are parallel, they have the same direction →−H, a hyperplane  
证据。图23.21说明了泰雷兹定理。我们画了一个证明，把细节留作练习。因为h1，h2，h3是平行的，它们有相同的方向→−h，一个超平面

### 23.9. AFFINE GEOMETRY: A GLIMPSE 23.9。仿射几何：一瞥

in →−E. Let be any nonnull vector such that A = a1+Ru. Since A is not parallel to H, we have E = H→−⊕Ru, and thus we can define the linear map p:→→−E → Ru, the projection on Ru parallel to H. This linear map induces an affine map f : E A, by defining f such that  
在→−e.中，设为任意非零矢量，使a=a1+ru。由于a不平行于h，我们有e=h→−ru，因此我们可以定义线性映射p：→→−e→ru，即ru上平行于h的投影。该线性映射通过定义f来诱导仿射映射f:e a，从而

f(b1 + w) = a1 + p(w),  
f（b1+w）=a1+p（w）

for all w ∈ →−E. Clearly, f(b1) = a1, and since H1,H2,H3 all have direction →−H, we also have f(b2) = a2 and f(b3) = a3. Since f is affine, it preserves ratios, and thus  
对于所有w∈→−e，很明显，f（b1）=a1，由于h1、h2、h3都有方向→−h，我们也有f（b2）=a2和f（b3）=a3。因为f是仿射的，所以它保持比率，因此

.  
.

The converse is immediate.   
反过来说是直接的。

We also have the following simple proposition, whose proof is left as an easy exercise.  
我们也有以下简单的命题，它的证明是一个简单的练习。

Proposition 23.11. Given any affine space E, given any two distinct points a,b ∈ E, and for any affine dilatation f different from the identity, if a0 = f(a), D = ha,bi is the line passing through a and b, and D0 is the line parallel to D and passing through a0, the following are equivalent:  
提案23.11.给定任意仿射空间e，给定任意两个不同点a，b∈e，并且对于任何不同于恒等式的仿射展开f，如果a0=f（a），d=ha，bi是通过a和b的线，而d0是平行于d并通过a0的线，则以下是等价的：

1. b0 = f(b);  
   b0=f（b）；
2. If f is a translation, then b0 is the intersection of D0 with the line parallel to ha,a0i passing through b;  
   如果f是平移，则b0是d0与平行于ha，a0i通过b的线的交点；

If f is a dilatation of center c, then b0 = D0 ∩ hc,bi.  
如果f是中心c的膨胀，则b0=d0 hc，bi。

The first case is the parallelogram law, and the second case follows easily from Thales’ theorem. For an illustration, see Figure 23.22.  
第一种情况是平行四边形定律，第二种情况很容易从泰雷兹定理得出。有关说明，请参见图23.22。

We are now ready to prove two classical results of affine geometry, Pappus’s theorem and Desargues’s theorem. Actually, these results are theorems of projective geometry, and we are stating affine versions of these important results. There are stronger versions that are best proved using projective geometry.  
我们现在准备证明仿射几何的两个经典结果，帕普斯定理和德沙格定理。实际上，这些结果是射影几何的定理，我们正在陈述这些重要结果的仿射形式。有更强大的版本，最好证明使用射影几何。

Proposition 23.12. Given any affine plane E, any two distinct lines D and D0, then for any distinct points a,b,c on D and a0,b0,c0 on D0, if a,b,c,a0, b0, c0 are distinct from the intersection of D and D0 (if D and D0 intersect) and if the lines ha,b0i and ha0,bi are parallel, and the lines hb,c0i and hb0,ci are parallel, then the lines ha,c0i and ha0,ci are parallel.  
提案23.12。给定任意仿射平面e，任意两条不同的线d和d0，那么对于任意不同的点a，b，c，d和a0，b0，c0，d0，如果a，b，c，a0，b0，c0与d和d0的交点不同（如果d和d0相交），并且如果线ha，b0i和ha0，bi平行，并且线hb，c0i和hb0，ci是para那么线ha，c0i和ha0，ci是平行的。

a

b

D

a’ = f(a)

D’

a

b

D

a’ = f(a)

D’

b’ = f(b)

a

b

D

a’ = f(a)

D’

c

b’ = f(b)

Figure 23.22: An illustration of Proposition 23.11. The bottom left diagram illustrates a translation, while the bottom right illustrates a central dilation through c.  
图23.22：提案23.11的说明。左下角的图表说明了平移，而右下角的图表说明了通过C的中心扩张。

Proof. Pappus’s theorem is illustrated in Figure 23.23. If D and D0 are not parallel, let d be their intersection. Let f be the dilatation of center d such that f(a) = b, and let g be the dilatation of center d such that g(b) = c. Since the lines ha,b0i anda0ha=0,bfi(bare parallel, and0) and b0 = g(c0).  
证据。Pappus定理如图23.23所示。如果d和d0不平行，则d是它们的交点。设f为中心d的扩张，使f（a）=b，设g为中心d的扩张，使g（b）=c。由于线ha，b0i和0ha=0，bfi（裸平行，and0）和b0=g（c0）。

However, we observed that dilatations with the same center commute, and thusand thus, lettingthe lines hb,c0i andh =hb0g,c◦ifare parallel, by Proposition 23.11 we have, we get c =Dh(aand) andD0aare parallel, we use translations instead of0 = h(c0). Again, by Proposition 23.11, thef ◦g = g◦f, lines ha,c0i and ha0,ci are parallel. If dilatations.   
然而，我们观察到在相同的中心通勤条件下的扩张，因此，假设线hb，c0i和h=hb0g，c如果是平行的，根据23.11号命题，我们得到c=dh（aand）和d0aare平行，我们使用翻译而不是0=h（c0）。同样，根据23.11号提案，off g=g f，行ha，c0i和ha0，ci是平行的。如果膨胀。

There is a converse to Pappus’s theorem, which yields a fancier version of Pappus’s theorem, but it is easier to prove it using projective geometry. It should be noted that in axiomatic presentations of projective geometry, Pappus’s theorem is equivalent to the commutativity of the ground field K (in the present case, K = R). We now prove an affine version of Desargues’s theorem.  
有一个与Pappus定理相反的定理，它产生了Pappus定理的更高级的版本，但使用射影几何更容易证明它。应该注意的是，在射影几何的公理化表示中，Pappus定理等价于基场k的交换性（在本例中，k=r）。我们现在证明了德沙格定理的仿射形式。

Proposition 23.13. Given any affine space E, and given any two triangles (a,b,c) and (a0,b0,c0), where a,b,c,a0,b0,c0 are all distinct, if ha,bi and ha0,b0i are parallel and hb,ci and hb0,c0i are parallel, then ha,ci and ha0,c0i are parallel iff the lines ha,a0i, hb,b0i, and hc,c0i are either parallel or concurrent (i.e., intersect in a common point).  
提案23.13。给定任意仿射空间e，给定任意两个三角形（a，b，c）和（a0，b0，c0），其中a，b，c，a0，b0，c0都是不同的，如果ha，bi和ha0，b0i是平行的，hb，ci和hb0，c0i是平行的，那么ha，ci和ha0，c0i是平行的，如果ha，a0i，hb，b0i和hc，c0i是平行的或同时的。（也就是说，在一个公共点上相交）。

hProof. We prove half of the proposition, the direction in which it is assumed that haa,c0,bi0iandare  
H车顶。我们证明了这个命题的一半，即假设haa，c0，bi0iandare的方向。

a0,c0i are parallel, leaving the converse as an exercise. Since the lines ha,bi and h  
a0，c0i是平行的，把相反的留作练习。因为行哈，bi和h

### 23.9. AFFINE GEOMETRY: A GLIMPSE 23.9。仿射几何：一瞥

*a*

*c*

*b*

*b*

*c*

*a*

*D*

*D*

Figure 23.23: Pappus’s theorem (affine version).  
图23.23:Pappus定理（仿射版本）。

parallel, the points a,b,a0,b0 are coplanar. Thus, either ha,a0i and hb,b0i are parallel, or they have some intersection d. We consider the second case where they intersect, leaving the other case as an easy exercise. Let f be the dilatation of center d such that f(a) = a0. By Proposition 23.11, we get f(b) = b0. If f(c) = c00, again by Proposition 23.11 twice, the  
平行，点A，B，A0，B0是共面的。因此，要么ha，a0i和hb，b0i是平行的，要么它们有一些交点d。我们考虑它们相交的第二种情况，将另一种情况留作一个简单的练习。设f为中心d的膨胀，使f（a）=a0。根据23.11号提案，我们得到f（b）=b0。如果f（c）=c00，再根据23.11号提案，两次

hfollows thatlines0 0hb,care parallel. Thus, the linesi andc00 =hb0c,c0. Indeed, recall that00i are parallel, and the lineshb,ci andhha,cb0,ci0iandare identical, and similarly the linesare parallel, and similarlyha0,c00i are parallel. From this itha,ci and a ,c i   
沿着这条线走，注意平行。因此，线si和c00=hb0c，c0。事实上，记得00i是平行的，线SHb，ci和hha，cb0，ci0i是相同的，同样的线是平行的，同样的，hhha，c0i是平行的。从这一点上来说，伊莎，CI和A，C I

ha0,c00i and ha0,c0i are identical. Since a0c0 and b0c0 are linearly independent, these lines have a unique intersection, which must be c00 = c0.  
HA0、C00i和HA0、C0i是相同的。由于a0c0和b0c0是线性无关的，所以这些线有一个唯一的交点，必须是c00=c0。

The direction where it is assumed that the lines ha,a0i, hb,b0i and hc,c0i, are either parallel or concurrent is left as an exercise (in fact, the proof is quite similar).   
假设线ha、a0i、hb、b0i和hc、c0i是平行的或同时存在的方向作为练习（事实上，证明是非常相似的）。

Desargues’s theorem is illustrated in Figure 23.24.  
德沙格定理如图23.24所示。

There is a fancier version of Desargues’s theorem, but it is easier to prove it using projective geometry. It should be noted that in axiomatic presentations of projective geometry, Desargues’s theorem is related to the associativity of the ground field K (in the present case, K = R). Also, Desargues’s theorem yields a geometric characterization of the affine dilatations. An affine dilatation f on an affine space E is a bijection that maps every line D to a line f(D) parallel to D. We leave the proof as an exercise.  
德沙格定理有一个更高级的版本，但是用射影几何来证明它更容易。应该注意的是，在射影几何的公理表示中，Desargues定理与地面场k的关联性有关（在本例中，k=r）。此外，德沙格定理给出了仿射扩张的几何特征。仿射空间e上的仿射展开式f是一个双射，它把每一条d线映射到平行于d的f（d）线上。我们把证明留作练习。

*d*

*a*

*b*

*c*

*a*

*b*

*c*

Figure 23.24: Desargues’s theorem (affine version).  
图23.24：德沙格定理（仿射版）。

## 23.10 Affine Hyperplanes 23.10仿射超平面

We now consider affine forms and affine hyperplanes. In Section 23.5 we observed that the set L of solutions of an equation  
我们现在考虑仿射形式和仿射超平面。在第23.5节中，我们观察到一个方程的解集l

ax + by = c  
ax+by=c

is an affine subspace of A2 of dimension 1, in fact, a line (provided that a and b are not both null). It would be equally easy to show that the set P of solutions of an equation  
是维度1的a2的仿射子空间，实际上是一行（前提是a和b都不是空的）。同样容易证明方程的解集p

ax + by + cz = d  
ax+by+cz=d

is an affine subspace of A3 of dimension 2, in fact, a plane (provided that a,b,c are not all null). More generally, the set H of solutions of an equation  
是维度2的a3的仿射子空间，实际上是一个平面（前提是a、b、c不是全部为空）。更一般地说，一个方程的解的集合h

λ1x1 + ··· + λmxm = µ  
λ1x1+····+λmxm=μ

is an affine subspace of Am, and if λ1,...,λm are not all null, it turns out that it is a subspace of dimension m − 1 called a hyperplane. We can interpret the equation  
是a m的仿射子空间，如果λ1，…，λm不都为空，则证明它是维度m-1的子空间，称为超平面。我们可以解释这个方程

λ1x1 + ··· + λmxm = µ  
λ1x1+····+λmxm=μ

in terms of the map f : Rm → R defined such that  
根据图f:rm→r，定义如下：

f(x1,...,xm) = λ1x1 + ··· + λmxm − µ  
f（x1，…，xm）=λ1x1+····+λmxm−礹

for all (x1,...,xm) ∈ Rm. It is immediately verified that this map is affine, and the set H of solutions of the equation  
对于所有（x1，…，xm）∈rm。立即证明该映射是仿射的，并且方程的解集h

λ1x1 + ··· + λmxm = µ  
λ1x1+····+λmxm=μ

23.10. AFFINE HYPERPLANES is the null set, or kernel, of the affine map f : Am → R, in the sense that  
第23.10条。仿射超平面是仿射映射f:am→r的空集或核，在这个意义上

H = f−1(0) = {x ∈ Am | f(x) = 0},  
h=f−1（0）=x∈am f（x）=0，

where x = (x1,...,xm).  
其中x=（x1，…，xm）。

Thus, it is interesting to consider affine forms, which are just affine maps f : E → R from an affine space to R. Unlike linear forms f∗, for which Kerf∗ is never empty (since it always contains the vector 0), it is possible that f−1(0) = ∅ for an affine form f. Given an affine map f : E → R, we also denote f−1(0) by Kerf, and we call it the kernel of f. Recall that an (affine) hyperplane is an affine subspace of codimension 1. The relationship between affine hyperplanes and affine forms is given by the following proposition.  
因此，考虑仿射形式是很有趣的，它只是仿射空间到r的仿射映射f:e→r。与线性形式f不同，对于这种形式，切口从不为空（因为它总是包含向量0），对于仿射形式f，f−1（0）=→r，我们也用切口表示f−1（0），我们称之为f的核心。回想一下，（仿射）超平面是余维1的仿射子空间。仿射超平面与仿射形式之间的关系由以下命题给出。

Proposition 23.14. Let E be an affine space. The following properties hold:  
提案23.14.设e为仿射空间。以下属性保留：

1. Given any nonconstant affine form f : E → R, its kernel H = Kerf is a hyperplane.  
   对于任何非常量仿射形式f:e→r，其核心h=kerf是超平面。
2. For any hyperplane H in E, there is a nonconstant affine form f : E → R such that H = Kerf. For any other affine form g: E → R such that H = Kerg, there is some λ ∈ R such that g = λf (with λ = 06 ).  
   对于e中的任何超平面h，都有一个非恒定的仿射形式f:e→r，这样h=kerf。对于任何其他的仿射形式g:e→r，例如h=kerg，有一些λ∈r，例如g=λf（带有λ=06）。
3. Given any hyperplane H in E and any (nonconstant) affine form f : E → R such that H = Kerf, every hyperplane H0 parallel to H is defined by a nonconstant affine form g such that g(a) = f(a) − λ, for all a ∈ E and some λ ∈ R.  
   给定e中的任意超平面h和任意（非常数）仿射形式f:e→r，使得h=kerf，每个平行于h的超平面h0都由非常数仿射形式g定义，这样g（a）=f（a）−λ，对于所有a∈e和一些λ∈r。

Proof. The proof is straightforward, and is omitted. It is also given in Gallier [71].   
证据。证据很直接，被省略了。加利尔文[71]也给出了这一点。

When E is of dimension n, given an affine frame (a0,(u1,...,un)) of E with origin a0, recall from Definition 23.5 that every point of E can be expressed uniquely as x = a0 +x1u1 +···+xnun, where (x1,...,xn) are the coordinates of x with respect to the affine frame (a0,(u1,...,un)).  
当e为n维数时，给定e的原点为a0的仿射帧（a0，（u1，…，un）），从定义23.5中回忆，e的每个点可以唯一地表示为x=a0+x1u1+······+xnun，其中（x1，…，xn）是x相对于仿射帧的坐标（a0，（u1，…，un））。

Also recall that every linear form f∗ is such that f∗(x) = λ1x1 + ··· + λnxn, for every x = x1u1 +···+xnun and some→− λ1,...,λn ∈ R. Since an affine form f : E → R satisfies the property f(a0 +x) = f(a0)+ f (x), denoting f(a0 +x) by f(x1,...,xn), we see that we have  
还记得，每一个线性形式f是这样的：f（x）=λ1x1+·····+λnxn，对于每一个x=x1u1+······+xnun和一些·−λ1，…，λn∈r。由于仿射形式f:e→r满足性质f（a0+x）=f（a0）+f（x），用f（x1，…，xn表示f（a0+x），我们看到我们已经

f(x1,...,xn) = λ1x1 + ··· + λnxn + µ,  
f（x1，…，xn）=λ1x1+····+λnxn+μ，

where µ = f(a0) ∈ R and λ1,...,λn ∈ R. Thus, a hyperplane is the set of points whose coordinates (x1,...,xn) satisfy the (affine) equation  
式中，μ=f（a0）∈r和λ1，…，λn∈r。因此，超平面是坐标（x1，…，xn）满足（仿射）方程的点集。

λ1x1 + ··· + λnxn + µ = 0.  
λ1x1+····+λnxn+μ=0.

## 23.11 Intersection of Affine Spaces 23.11仿射空间的交集

In this section we take a closer look at the intersection of affine subspaces. This subsection can be omitted at first reading.  
在本节中，我们将更详细地了解仿射子空间的交集。本小节可在第一次阅读时省略。

First, we need a result of linear algebra. Given a vector space E and any two subspaces M  
首先，我们需要一个线性代数的结果。给定向量空间e和任意两个子空间m

→ ⊕N and in2 : N → M⊕N  
→n和in2:n→m\_n

map fromof the inclusion map fromand thus, injectionsMand+NN, there are several interesting linear maps. We have the canonical injectionsandMj:∩NN→tofM:N+MNwith∩, the canonical injectionsMN ∩in2. Then, we have the mapsM N and g: M inN1 : MM f M+ g: M ∩ N → M ⊕iN: M, and→,  
包含图From的映射，因此，InjectionsMand+nn，有几个有趣的线性映射。我们有标准注射剂和mj：n n→tofm:n+mn和，标准注射剂mn in2。然后，我们有mapsm n和g:m inn1:mm f m+g:m n→m in:m，和→，

→N to⊕M with in1, and∩ →g is the composition of the inclusion⊕N, where f is the composition  
→n to m with in1，and→g is the composition of the inclusion n，其中f is the composition

i − j: M ⊕ N → M + N.  
i−j:m n→m+n。

Proposition 23.15. Given a vector space E and any two subspaces M and N, with the definitions above,  
提案23.15。给定一个向量空间e和任意两个子空间m和n，以及上面的定义，

0 −→ M ∩ N −f+→g M ⊕ N −i−→j M + N −→ 0  
0−→M N−F+→G M N−I−→J M+N−→0

Im(is a short exact sequence, which means thatf + g) = Ker(i − j). As a consequence, we have the Grassmann relationf + g is injective, i − j is surjective, and that  
im（是一个短的精确序列，这意味着f+g）=ker（i-j）。因此，我们得到格拉斯曼关系式+G是内射的，I−J是外射的，并且

dim(M) + dim(N) = dim(M + N) + dim(M ∩ N).  
尺寸（m）+尺寸（n）=dim（m+n）+尺寸（m n）。

Proof. It is obvious that i −, andj is surjective and thatv ∈ Ni. Then,(u) = j(iv() =u) =wfj∈(+vj)M)g, and thus, by definition of, as desired. The second part ofis injective. Assume that (∩ N. By definition of f andi andi −g,  
证据。很明显，i−和j是主观性的，v∈ni。然后，（u）=j（iv（）=u）=wfj∈（+vj）m）g，因此，根据定义，根据需要。第二部分是注射剂。假设（n.根据f和i和i-g的定义，

j)(u + v) = 0, where u ∈ MN, such that j, there is some w ∈g(Mw)∩, and thus Im(f + g) = Ker(i − u = f(w) and v =  
j）（u+v）=0，其中u∈mn，这样j，有一些w∈g（mw），因此im（f+g）=ker（i−u=f（w）和v=

the proposition follows from standard results of linear algebra (see Artin [7], Strang [165], or Lang [106]).   
这个命题来自线性代数的标准结果（见Artin[7]、Strang[165]或Lang[106]）。

We now prove a simple proposition about the intersection of affine subspaces.  
我们现在证明一个关于仿射子空间交集的简单命题。

Proposition 23.16. Given any affine space E, for any two nonempty affine subspaces M and N, the following facts hold:  
提案23.16。对于任意两个非空的仿射子空间m和n，对于任意仿射空间e，以下事实成立：

1. M ∩ N 6= ∅ iff →−ab ∈ −M→ + →−N for some a ∈ M and some b ∈ N.  
   m n 6=∅iff→−a b∈−−m→＋→−n，对于某些a∈m和一些b∈n。
2. M ∩−→N∩consists of a single point iff→−N = {0}. →−ab ∈ M−→ + →−N for some a ∈ M and some b ∈ N, and M  
   M−→N由单点iff→−N=0组成。对于一些a∈m和一些b∈n，和m
3. If S is the least affine subspace containing M and N, then →−S = −M→ + →−N + K→−ab (the vector space →−E is defined over the field K).  
   如果s是包含m和n的最小仿射子空间，则→−s=−m→+→−n+k→−ab（矢量空间→−e在字段k上定义）。

### 23.11. INTERSECTION OF AFFINE SPACES 11月23日。仿射空间的交集

Proof.−→(1) Pick any a ∈ M and any→− b→−∈ N, which is possible, since M and N are nonempty.we have  
证明。−→（1）选择任意a∈m和任意→−b→−n，这是可能的，因为m和n是非空的。

→− , withfor some→−ac ∈a M∈ Mandand somebc ∈ N, and thus,b ∈ N. Thenab ∈→−abM=+−ax→N+. Conversely, assume that→−by, for some x ∈ M and  
→−，对于一些→−ac∈a m∈m and和一些bc∈n，因此，b∈n.thenab∈→−abm=+−ax→n+。相反，假设→−by，对于某些x∈m和

some y ∈ N. But we also have →−ab = ax−→ + −xy→ + →−yb,  
一些y∈n，但我们也有→−ab=ax−→+xy→+yb，

[x,y], and since −yx→ = 2  
[x，y]，并且由于−yx→=2

and thus we get 0 =(y,−1). Thus x also belongs tox−xy→∈ M→−yb+,→−∩ybxN−= 2, and→−bybN, that is,−, sinceMy∩is the barycenter of the weighted points (N−xy→N6= = 2∅being an affine subspace, it is closed under. →−by. Thus, b is the middle of the segmentb,2) and  
因此我们得到0=（Y，−1）。因此x也属于tox−xy→∈m→−yb+、→yb x n−2，和→−bybn，即−，sincemy是加权点的重心（n−xy→n6＝2∅是仿射子空间，它在下面闭合。→−通过。因此，b是节段b，2）和

barycenters. Thus,  
重心。因此，

1. Note that in general, if M ∩ N 6= ∅, then  
   注意，一般情况下，如果m n 6=∅，则

−−−−M ∩→N = −M→ ∩ →−N,  
−−−M→N=−M→→−N，

because  
因为

−−−−M ∩→N = {→−ab | a,b ∈ M ∩ N} = {→−ab | a,b ∈ M} ∩ {→−ab | a,b ∈ N} = −M→ ∩ →−N.  
−−−m→n=→−a b a，b∈m n→−ab a，b∈m→−ab a，b∈n=−m→→−n。

Since M ∩ N = c + −−−−M ∩→N for any c ∈ M ∩ N, we have  
因为m n=c+−−−−m→n对于任何c∈m n，我们有

M ∩ N = c + −M→ ∩ →−N for any c ∈ M ∩ N.  
m n=c+−m−n表示任意c m n。

This fact together with what we proved in (1) proves (2).From this it follows that if M∩N 6= ∅, then M∩N consists of a single point iff −M→∩→−N = {0}.  
这一事实连同我们在（1）中证明的事实证明（2），由此得出，如果m n 6=∅，那么m n由单点iff−m n 0组成。

1. This is left as an easy exercise.   
   这是一个简单的练习。

Remarks:  
评论：

1. aThe proof of Proposition 23.16 shows that if∈ M and all b ∈ N. M ∩ N =6 ∅, then →−ab ∈ −M→ + →−N for all  
   A.命题23.16的证明表明，如果∈m和所有b∈n.m n=6∅，那么→−ab∈−m→＋→−n代表所有
2. Proposition 23.16 implies that for any two nonempty affine subspaces , if  
   命题23.16意味着对于任意两个非空仿射子空间，如果

, then M ∩ N consists of a single point. Indeed, if E = M  
，则m n由一个点组成。实际上，如果e=m

part (2) of the proposition.ab ∈ E for all a ∈ M and all b ∈ N, and since M−→ ∩ →−N = {0}, the result follows from⊕ N, then  
命题的第（2）部分，a b∈e表示所有a∈m和所有b∈n，由于m−→→−n=0，其结果如下n，则

We can now state the following proposition.  
我们现在可以陈述以下命题。

Proposition 23.17. Given an affine space E and any two nonempty affine subspaces M and N, if S is the least affine subspace containing M and N, then the following properties hold:  
提案23.17。给定一个仿射空间e和任意两个非空仿射子空间m和n，如果s是包含m和n的最小仿射子空间，则以下属性成立：

1. If M ∩ N = ∅, then  
   如果m n=直径，则

dim(M) + dim(N) < dim(E) + dim(−M→ + →−N)  
尺寸（m）+尺寸（n）<尺寸（e）+尺寸（m→+N）

and dim(S) = dim(M) + dim(N) + 1 − dim(−M→ ∩ →−N).  
和dim（s）=dim（m）+dim（n）+1−dim（−m→→−n）。

1. If M ∩ N 6= ∅, then  
   如果m n 6=直径，则

dim(S) = dim(M) + dim(N) − dim(M ∩ N).  
尺寸=尺寸（m）+尺寸（n）-尺寸（m n）。

Proof. The proof is not difficult, using Proposition 23.16 and Proposition 23.15, but we leave it as an exercise.   
证据。使用23.16号和23.15号提案，证明并不困难，但我们将其作为练习。

Chapter 24  
第二十四章

# Embedding an Affine Space in a Vector Space 在向量空间中嵌入仿射空间

## 24.1 The “Hat Construction,” or Homogenizing 24.1“帽结构”或均质化

For all practical purposes, most geometric objects, including curves and surfaces, live in affine spaces. A disadvantage of the affine world is that points and vectors live in disjoint universes. It is often more convenient, at least mathematically, to deal with linear objects (vector spaces, linear combinations, linear maps), rather than affine objects (affine spaces, affine combinations, affine maps). Actually, it would also be advantageous if we could manipulate points and vectors as if they lived in a common universe, using perhaps an extra bit of information to distinguish between them if necessary.  
在所有实际应用中，大多数几何对象，包括曲线和曲面，都生活在仿射空间中。仿射世界的一个缺点是点和向量生活在不相交的宇宙中。至少在数学上，处理线性对象（向量空间、线性组合、线性映射）比处理仿射对象（仿射空间、仿射组合、仿射映射）更方便。事实上，如果我们能像生活在一个共同的宇宙中一样操纵点和向量，在必要的时候使用额外的信息来区分它们，这也是有利的。

Such a “homogenization” (or “hat construction”) can be achieved. As a matter of fact, such a homogenization of an affine space and its associated vector space will be very useful to define and manipulate rational curves and surfaces. Indeed, the hat construction yields a canonical construction of the projective completion of an affine space. It also leads to a very elegant method for obtaining the various formulae giving the derivatives of a polynomial curve, or the directional derivatives of polynomial surfaces. However, these formulae are not needed here. Thus we omit this topic, referring the readers to Gallier [71].  
这样的“同质化”（或“帽子结构”）可以实现。实际上，仿射空间及其相关向量空间的这种均匀化对于定义和操纵有理曲线和曲面是非常有用的。实际上，hat构造生成仿射空间射影完备的规范构造。它还引出了一种非常优雅的方法，用于获得给出多项式曲线导数或多项式曲面方向导数的各种公式。然而，这里不需要这些公式。因此，我们省略了这一主题，将读者引向Gallier[71]。

This chapter proceeds as follows. First, the construction of a vector space Eb in which both E and →−E are embedded as (affine) hyperplanes is described. It is shown how affine frames in E become bases in Eb. It turns out that Eb is characterized by a universality property: Affine maps to vector spaces extend uniquely to linear maps. As a consequence, affine maps between affine spaces E and F extend to linear maps between Eb and Fb.  
本章内容如下。首先，描述了向量空间eb的构造，其中e和→−e都嵌入（仿射）超平面。它显示了e中的仿射帧如何成为eb中的基。结果表明，电子束具有普适性特征：矢量空间的仿射映射唯一地扩展到线性映射。因此，仿射空间e和f之间的仿射映射扩展到eb和fb之间的线性映射。

Let us first explain how to distinguish between points and vectors practically, using what amounts to a “hacking trick.” Then, we will show that such a procedure can be put on firm mathematical grounds.  
让我们先解释一下如何在实际中区分点和向量，使用相当于“黑客把戏”的方法。然后，我们将展示这样一个过程可以建立在坚实的数学基础上。

Assume that we consider the real affine space E of dimension 3, and that we have some  
假设我们考虑维3的实仿射空间e，我们有一些

741  
七百四十一

affine frame (a0,(v1,v2,v2)). With respect to this affine frame, every point x ∈ E is repre-→− sented by its coordinates (x1,x2,x3), where a = a0 + x1v1 + x2v2 + x3v3. A vector u ∈ E is also represented by its coordinates (u1,u2,u3) over the basis (v1,v2,v2). One way to distinguish between points and vectors is to add a fourth coordinate, and to agree that points are represented by (row) vectors (x1,x2,x3,1) whose fourth coordinate is 1, and that vectors are represented by (row) vectors (v1,v2,v3,0) whose fourth coordinate is 0. This “programming trick” actually works very well. Of course, we are opening the door for strange elements such as (x1,x2,x3,5), where the fourth coordinate is neither 1 nor 0.  
仿射帧（a0，（v1，v2，v2））。对于这个仿射框架，每个点x∈e由其坐标（x1，x2，x3）表示，其中a=a0+x1v1+x2v2+x3v3。矢量u∈e也由其在基（v1，v2，v2）上的坐标（u1，u2，u3）表示。区分点和矢量的一种方法是添加第四个坐标，并同意点由第四个坐标为1的（行）矢量（x1，x2，x3,1）表示，矢量由第四个坐标为0的（行）矢量（v1，v2，v3,0）表示。这个“编程技巧”实际上非常有效。当然，我们打开了一些奇怪元素的门，比如（x1，x2，x3，5），其中第四个坐标既不是1也不是0。

The question is, can we make sense of such elements, and of such a construction? The answer is yes. We will present a construction in which an affine space E, is embedded in a vector space Eb, in which →−E is embedded as a hyperplane passing through the origin, and E itself is embedded as an affine hyperplane, defined as ω−1(1), for some linear form ω: Eb → R. In the case of an affine space→− E of dimension 2, we can think of Eb as the vector space R3 of dimension 3 in which E corresponds to the xy-plane, and E corresponds to the plane of equation z = 1, parallel to the xy-plane and passing through the point on the z-axis of coordinates (0,0,1). The construction of the vector space Eb is presented in some detail in Berger [11]. Berger explains the construction in terms of vector fields. We prefer a more geometric and simpler description in terms of simple geometric transformations, translations, and dilatations.  
问题是，我们能理解这样的元素和这样的结构吗？答案是肯定的。我们将提出一种构造，其中仿射空间e嵌入到向量空间eb中，其中→−e嵌入为穿过原点的超平面，e本身嵌入为仿射超平面，定义为ω−1（1），对于某种线性形式ω：eb→r。对于仿射空间，则定义为ω−1（1）。速度→−e在尺寸2中，我们可以把eb看作是尺寸3的向量空间r3，其中e对应于xy平面，e对应于方程z=1的平面，平行于xy平面，并通过坐标z轴上的点（0,0,1）。向量空间eb的构造在Berger[11]中有详细介绍。Berger解释了矢量场的构造。我们更喜欢用简单的几何变换、翻译和扩张来描述更为几何和简单的描述。

Remark: Readers with a good knowledge of geometry will recognize the first step in embedding an affine space into a projective space. We will also show that the homogenization Eb of an affine space E,, satisfies a universal property with respect to the extension of affine maps to linear maps. As a consequence, the vector space Eb is unique up to isomorphism, and its actual construction is not so important. However, it is quite useful to visualize the space Eb, in order to understand well rational curves and rational surfaces.  
注：具有良好几何知识的读者将认识到在射影空间中嵌入仿射空间的第一步。我们还将证明仿射空间e，，的均匀化eb满足仿射映射到线性映射扩展的一个普遍性质。因此，矢量空间eb具有独特的同构性，其实际构造并不重要。然而，为了更好地理解有理曲线和有理曲面，可视化空间eb是非常有用的。

As usual, for simplicity, it is assumed that all vector spaces are defined over the field R of real numbers, and that all families of scalars (points and vectors) are finite. The extension to arbitrary fields and to families of finite support is immediate. We begin by defining two very simple kinds of geometric (affine) transformations. Given an affine space E,, every u ∈ →−E induces a mapping tu : E → E, called a translation, and defined such that for every a ∈ E. Clearly, the set of translations is a vector space isomorphic to  
通常，为了简单起见，假设所有向量空间都定义在实数的r域上，并且所有标量族（点和向量）都是有限的。对任意域和有限支撑族的扩展是直接的。我们首先定义两种非常简单的几何（仿射）变换。给定一个仿射空间e，每一个u∈→−e诱导一个映射tu:e→e，称为翻译，并定义为每一个a∈e。显然，翻译集是一个向量空间同构于

E. Thus, we will use the same notation u for both the vector u and the translation tu. Given any point a and any scalar λ ∈ R, we define the mapping Ha,λ : E → E, called dilatation (or central dilatation, or homothety) of center a and ratio λ, and defined such that  
e.因此，我们将对向量u和平移tu使用相同的符号u。给定任何点a和任何标量λ∈r，我们定义映射ha，λ：e→e，称为中心a和比率λ的扩张（或中心扩张，或同构），并定义如下：

Ha,λ(x) = a + λax,−→  
ha，λ（x）=a+λax，−→

for every x ∈ E. We have Ha,λ(a) = a, and when−→λ = 06 and x 6= a, Ha,λ(x) is on the line defined by a and x, and is obtained by “scaling” ax by λ. The effect is a uniform dilatation (or contraction, if λ < 1). When λ = 0, Ha,0(x) = a for all x ∈ E, and Ha,0 is the constant affine map sending every point to a. If we assume λ = 16 , note that Ha,λ is never the identity, and since a is a fixed point, Ha,λ is never a translation.  
对于每一个x∈e，我们有ha，λ（a）=a，当−→λ=06和x 6=a，ha，λ（x）在a和x定义的线上时，由λ“缩放”ax得到。效果是均匀的扩张（或收缩，如果λ<1）。当λ=0，ha，0（x）=a时，对于所有x∈e，ha，0是将每个点发送到a的常数仿射映射。如果假设λ=16，请注意，ha，λ绝不是同一性，并且由于a是固定点，ha，λ绝不是翻译。

We now consider the set Eb of geometric transformations from E to E, consisting of the union of the (disjoint) sets of translations and dilatations of ratio λ = 16 . We would like→− to give this set the structure of a vector space, in such a way that both E and E can be naturally embedded into Eb. In fact, it will turn out that barycenters show up quite naturally too!  
我们现在考虑从e到e的几何变换集eb，它由（不相交的）平移集的并集和比率λ=16的扩张组成。我们希望→−给这个集合一个向量空间的结构，这样E和E都可以自然地嵌入到eb中。事实上，重心也会很自然地出现！

In order to “add” two dilatations Ha1,λ1 and Ha2,λ2, it turns out that it is more convenient to consider dilatations of the form Ha,1−λ, where λ = 06 . To see this, let us see the effect of such a dilatation on a point x ∈ E: We have  
为了“加上”两个展开式Ha1，λ1和Ha2，λ2，结果表明考虑形式Ha，1−λ的展开更为方便，其中λ=06。为了看到这一点，让我们看看这种扩张对点x∈e的影响：我们有

Ha,1−λ(x) = a + (1 − λ)ax−→ = a + −ax→ − λ−ax→ = x + λxa.−→  
ha，1−λ（x）=a+（1−λ）ax−→=a+−ax→−λ−ax→=x+λxa.−→

For simplicity of notation, let us denote Ha,1−λ by ha,λi. Then, we have  
为了简化记法，让我们用ha，λi表示ha，1−λ。然后，我们得到

ha,λi(x) = x + λxa.−→  
ha，λi（x）=x+λxa.−→

Remarks:  
评论：

1. Note that Ha,1−λ(x) = Hx,λ(a).  
   注意，ha，1−λ（x）=hx，λ（a）。
2. Berger defines a map h: E → →−E as a vector field. Thus, each ha,λi can be viewed as the vector field x 7→ λxa−→. Similarly, a translation u can be viewed as the constant vector field x 7→ u. Thus, we could define Eb as the (disjoint) union of these two vector fields. We prefer our view in terms of geometric transformations. Then, since  
   Berger将地图h:e→→−e定义为矢量场。因此，可以将每个ha，λi视为向量场x 7→λxa−→。同样，一个平移u可以被看作是一个常数向量场x 7→u。因此，我们可以将eb定义为这两个向量场的（不相交）并集。我们更喜欢几何变换的观点。那么，自从

and ,  
而且，

if we want to define ha1,λ1i+b ha2,λ2i, we see that we have to distinguish between two cases:  
如果我们要定义ha1，λ1i+b ha2，λ2i，我们必须区分两种情况：

1. λ1 + λ2 = 0. In this case, since  
   λ1+λ2=0.在这种情况下，因为

λ1xa−→1 + λ2xa−→2 = λ1xa−→1 − λ1xa−→2 = λ1a−−2→a1,  
λ1xa−→1+λ2xa−→2=λ1xa−→1-λ1xa−→2=λ1a−2→a1，

we let  
我们让

ha1,λ1i +b ha2,λ2i = λ1a−−2→a1,  
ha1，λ1i+b ha2，λ2i=λ1a−2→a1，

where denotes the translation associated with the vector.  
其中表示与向量关联的转换。

1. λ1 +λ2 = 06 . In this case, the points a1 and a2 assigned the weights λ1/(λ1 +λ2) and λ2/(λ1 + λ2) have a barycenter  
   λ1+λ2=06.在这种情况下，分配给权重λ1/（λ1+λ2）和λ2/（λ1+λ2）的点a1和a2具有重心。

,  
，

such that  
这样的话

Since   
自从

we let  
我们让

,  
，

the dilatation associated with the point b and the scalar λ1 + λ2.  
与点B和标量λ1+λ2相关的膨胀。

Given a translation defined by u and a dilatation ha,λi, since λ = 06 , we have  
给定u定义的平移和扩张ha，λi，因为λ=06，我们得到

λxa−→ + u = λ(xa−→ + λ−1u),  
λxa−→+u=λ（xa−→+λ−1u）

and so, letting b = a + λ−1u, since →−ab = λ−1u, we have  
因此，假设b=a+λ−1u，因为→−ab=λ−1u，我们有

λxa−→ + u = λ(xa−→ + λ−1u) = λ(−xa→ + →−ab) = λ→−xb,  
λxa−→+u=λ（xa−→+λ−1u）=λ（−xa→+ab）=λ→−xb，

and we let ha,λi +b u = ha + λ−1u,λi,  
我们让ha，λi+b u=ha+λ−1u，λi，

the dilatation of center a + λ−1u and ratio λ.  
中心A+λ−1u和比率λ的扩张。

The sum of two translations u and v is of course defined as the translation u + v. It is also natural to define multiplication by a scalar as follows:  
两个翻译u和v的和当然定义为翻译u+v。自然地，定义一个标量的乘法如下：

µ · ha,λi = ha,λµi,  
礹·ha，λi=ha，λ礹i，

and  
和

where λu is the product by a scalar in  
式中，λu是中标量的乘积

We can now use the definition of the above operations to state the following proposition, showing that the “hat construction” described above has allowed us to achieve our goal of embedding both E and →−E in the vector space Eb.  
我们现在可以使用上述操作的定义来陈述以下命题，表明上述“帽结构”允许我们实现将e和→−e嵌入向量空间eb的目标。

Proposition 24.1. The set Eb consisting of the disjoint union of the translations and the dilatations Ha,1−λ = ha,λi, λ ∈ R,λ = 06 , is a vector space under the following operations of addition and multiplication by a scalar: If λ1 + λ2 = 0, then  
提案24.1.由平移与扩张的不相交并合ha，1−λ=ha，λi，λ∈r，λ=06组成的集合eb是一个在下列标量加乘运算下的向量空间：如果λ1+λ2=0，那么

;  
；

if λ1 + λ2 = 06 , then  
如果λ1+λ2=06，则

,  
，

u +b v = u + v;  
U+B V=U+V；

if µ 6= 0, then  
如果μ6=0，则

µ · ha,λi = ha,λµi, 0 · ha,λi = 0;  
μ·ha，λi=ha，λμi，0·ha，λi=0；

and  
和

λ · u = λu.  
λ·u=λu。

Furthermore, the map ω: Eb → R defined such that  
此外，图ω：eb→r定义如下：

ω(ha,λi) = λ, ω(u) = 0,  
ω（ha，λi）=λ，ω（u）=0，

is a linear form, ω−1(0) is a hyperplane isomorphic to →−E under the injective linear map such that i(u) = tu (the translation associated with u), and ω−1(1) is an affine  
是线性形式，ω−1（0）是在注入线性映射下与→−e同构的超平面，因此i（u）=tu（与u相关的平移），ω−1（1）是仿射

hyperplane isomorphic towhere j(a) = ha,1i for everyE with directiona ∈ E. Finally, for everyi(→−E), under the injective affine mapa ∈ E, we have j: E → Eb,  
超平面同构Towhere j（a）=ha，1i，对于方向为∈e的Everye，最后，对于Everyi（→−e），在注入仿射映射a∈e下，我们得到j:e→eb，

.  
.

Proof. The verification that Eb is a vector space is straightforward. The linear map mapping a vector u to the translation defined by u is clearly an injection i: →−E → Eb embedding →−E as an hyperplane in Eb. It is also clear that ω is a linear form. Note that  
证据。验证eb是一个向量空间是很简单的。将向量u映射到由u定义的平移的线性映射显然是注入i：→−e→eb嵌入→−e作为eb中的超平面。很明显，ω是一个线性形式。注意

j(a + u) = ha + u,1i = ha,1i +b u,  
j（a+u）=ha+u，1i=ha，1i+b u，

where u stands for the translation associated with the vector u, and thus j is an affine injection with associated linear map i. Thus, ω−1(1) is indeed an affine hyperplane isomorphic to E with direction, under the map j: E → Eb. Finally, from the definition of +b , for every a ∈ E and every u ∈ →−E, since  
式中，u代表与向量u相关的平移，因此j是具有相关线性映射i的仿射注入。因此，ω−1（1）确实是在映射j:e→eb下与e同构的仿射超平面。最后，从+b的定义来看，对于每一个a∈e和每一个u∈→−e，因为

i(u) +b λ · j(a) = u +b ha,λi = ha + λ−1u,λi,  
i（u）+bλ·j（a）=u+b ha，λi=ha+λ−1u，λi，

whenarbitrary elementλ = 06 , we get any arbitraryhb,µi, µ 6= 0, by pickingv ∈ Ebλby picking= µ and uλ== 0µ→−ab. Thus,and u = v, and we get any  
当双元素λ=06时，我们通过选择v∈ebλ通过选择=μ和uλ==0μ→−ab得到任意任意的hb，μi，μ6=0，因此，和u=v，我们得到

,  
，

and since , we have  
从那以后，我们

,  
，

for every a ∈ E.   
对于每一个a∈e。

Ω

*i*

*E*

=

*ω*

−

1

(

0

)

*u*

*j*

(

*E*

)

=

*ω*

−

1

(

1

)

*a*

*,*

*λ*

*a*

*,*

1

=

*a*

Figure 24.1: Embedding an affine space E, into a vector space Eb.  
图24.1：在向量空间eb中嵌入仿射空间e。

Figure 24.1 illustrates the embedding of the affine space E into the vector space Eb, when E is an affine plane.  
图24.1说明了当e是仿射平面时，仿射空间e嵌入向量空间eb。

Note that Eb is isomorphic to →−E ∪ (E × R∗). Intuitively, we can think of Eb as a stack of parallel hyperplanes, one for each λ, a little bit like an infinite stack of very thin pancakes! There are two privileged pancakes: one corresponding to E, for λ = 1, and one corresponding to →−E, for λ = 0.  
请注意，eb与→−e（e×r）同构。直观地说，我们可以把eb看作一堆平行超平面，每一个λ一个，有点像一堆无限薄的薄煎饼！有两个特权煎饼：一个对应于e，对于λ=1，另一个对应于→−e，对于λ=0。

From now on, we will identify j(E) and E, and and →−E. We will also write λa instead of ha,λi, which we will call a weighted point, and write 1a just as a. When we want to be more precise, we may also write ha,1i as a. In particular, when we consider the homogenized version Ab of the affine space A associated with the field R considered as an  
从现在开始，我们将识别j（e）和e，以及→−e。我们还将写出λa而不是ha，λi，我们称之为加权点，并将1a写成a。当我们想要更精确的时候，我们也可以把ha，1i写成a。特别是当我们考虑仿射的同质化版本ab时与字段r关联的空间a被视为

affine space, we write λ for hλ,1i, when viewing λ as a point in both A and Ab, and simply λ, when viewing λ as a vector in R and in Ab. As an example, the expression 2 + 3 denotes the real number 5, in A, (2 + 3)/2 denotes the midpoint of the segment , which can be denoted by 2.5, and 2+3 does not make sense in A, since it is not a barycentric combination. However, in Ab, the expression 2 + 3 makes sense: It is the weighted point. Then, in view of the fact that  
在仿射空间中，当把λ视为a和ab中的一个点时，我们为hλ，1i写出λ，当把λ视为r和ab中的一个矢量时，我们只写出λ。例如，表达式2+3表示实数5，在a中，（2+3）/2表示段的中点，可以用2.5和2+3表示。在A中没有意义，因为它不是重心组合。然而，在ab中，表达式2+3是有意义的：它是加权点。那么，鉴于

ha + u,1i = ha,1i +b u,  
ha+u，1i=ha，1i+b u，

and since we are identifying a + u with ha + u,1i (under the injection j), in the simplified notation the above reads as a + u = a +b u. Thus, we go one step further, and denote a +b u by a + u. However, since ha,λi +b u = ha + λ−1u,λi,  
由于我们将a+u与ha+u，1i（在注入j下）进行识别，在简化的符号中，上述内容读作a+u=a+b u。因此，我们更进一步，用a+u表示a+b u。但是，由于ha，λi+b u=ha+λ−1u，λi，

we will refrain from writing λa +b u as λa + u, because we find it too confusing. From Proposition 24.1, for every a ∈ E, every element of Eb can be written uniquely as u +b λa. We also denote λa + (b −µ)b  
我们将避免将λa+b u写成λa+u，因为我们发现它太令人困惑了。从命题24.1可以看出，对于每一个a∈e，e b的每一个元素都可以唯一地写成u+bλa。我们也表示λa+（b−μ）b。

by λa −b µb.  
通过λa−bμb。

We can now justify rigorously the programming trick of the introduction of an extra coordinate to distinguish between points and vectors. First, we make a few observations. Given any family (ai)i∈I of points in E, and any family (λi)i∈I of scalars in R, it is easily shown by induction on the size of I that the following holds:  
我们现在可以严格证明引入额外坐标来区分点和向量的编程技巧。首先，我们做了一些观察。给定e中点的任一族（ai）i∈i，r中标量的任一族（λi）i∈i，通过对i大小的归纳，可以很容易地表示如下：

(1) If Pi∈I λi = 0, then  
（1）如果pi∈iλi=0，则

,  
，

where  
哪里

X−−−−−→ X −→ λiai = λibai  
x−−−→x−→λiai=λibai

i∈I i∈I  
I∈I I∈I

for any b ∈ E, which, by Proposition 23.1, is a vector independent of b, or (2) If Pi∈I λi = 06 , then  
对于任何b∈e，根据23.1，它是独立于b的向量，或者（2）如果pi∈iλi=06，那么

.  
.

Thus, we see how barycenters reenter the scene quite naturally, and that in Eb, we can make sense ofh i Pi∈Ihai,λii, regardless of the value of−1(1), and thus it is a point. WhenPi∈I λi. When Pi∈I λi = 1P, the elementi∈I λi = 0,  
因此，我们可以很自然地看到重心是如何重新进入场景的，而在eb中，我们可以理解h i pi∈ihai，λii，而不考虑−1（1）的值，因此它是一个点。when pi∈iλi.当pi∈iλi=1p时，元素i∈iλi=0，

P ai,λi belongs to the hyperplane ω  
p ai，λi属于超平面ω

the linear combination of pointsi∈I Pi∈ λiai is a vector, and when I = {1,...,n}, we allow  
点si∈i pi∈λiai的线性组合是一个向量，当i=1，…，n时，我们允许

I  
我

ourselves to write  
我们自己写

λ1a1 +b ··· +b λnan,  
λ1a1+b····+bλnan，

where some of the occurrences of + can be replaced by , as  
其中一些出现的+可以替换为，例如

λ1a1 + ··· + λnan,  
λ1a1+····+λnan，

where the occurrences of −b (if any) are replaced by −.  
其中−b（如果有）的出现被−替换。

In fact, we have the following slightly more general property, which is left as an exercise.  
事实上，我们有以下稍微更一般的属性，作为练习。

Proposition 24.2. Given any affine space E,, for any family (ai)i∈I of points in E, any family (λi)i∈I of scalars in R, and any family (vj)j∈J of vectors in →−E, with I ∩ J = ∅, the following properties hold:  
提案24.2.给定任意仿射空间e，，对于任意族（ai）i∈i中的点，任意族（λi）i∈i中的标度r，任意族（vj）j∈j中的向量→−e，具有i j∅的性质如下：

1. If Pi∈I λi = 0, then  
   如果pi∈iλi=0，则

,  
，

where  
哪里

X−−−−−→ X −→ λiai = λibai  
x−−−→x−→λiai=λibai

i∈I i∈I  
I∈I I∈I

for any b ∈ E, which, by Proposition 23.1, is a vector independent of b, or  
对于任何b∈e，根据23.1，它是独立于b的向量，或

1. If Pi∈I λi = 06 , then  
   如果pi∈iλi=06，则

.  
.

Proof. By induction on the size of I and the size of J.   
证据。通过归纳I的大小和J的大小。

The above formulae show that we have some kind of extended barycentric calculus. Operations on weighted points and vectors were introduced by H. Grassmann, in his book published in 1844! This calculus will be helpful in dealing with rational curves.  
上面的公式表明我们有某种扩展的重心微积分。在1844年出版的书中，H.Grassmann介绍了加权点和向量的运算。这种计算方法将有助于处理有理曲线。

## 24.2 Affine Frames of E and Bases of Eb 24.2 e的仿射框架和eb的基

There is also a nice relationship between affine frames in E, and bases of Eb, stated in the following proposition.  
在e中的仿射框架和eb的基之间也有一个很好的关系，如下所述。

Proposition 24.3. Given any affine space E,, for any affine frame , for E, the family is a basis for Eb, and for any affine frame  
提案24.3.对于任意仿射空间e，对于任意仿射框架，对于e，族是eb和任意仿射框架的基础。

(a0,...,am) for E, the family (a0,...,am) is a basis for Eb. Furthermore, given any element hx,λi ∈ Eb, if  
（a0，…，am）对于e，家庭（a0，…，am）是eb的基础。此外，给定任何元素hx，λi∈eb，如果



over the affine frame , then the coordinates of hx,λi over the basis  
在仿射框架上，然后是Hx的坐标，在基上的λi

are  
是

(λx1,...,λxm,λ).  
（λx1，…，λxm，λ）。

For any vector v ∈ →−E, if  
对于任何向量v∈→−e，如果



### 24.2. AFFINE FRAMES OF E AND BASES OF Eˆ 24.2。e的仿射框架和e\_的基

over the basis , then over the basis , the coordinates of v are  
在基础上，然后在基础上，v的坐标是

(v1,...,vm,0).  
（v1，…，vm，0）。

For any element ha,λi, where λ = 06 , if the barycentric coordinates of a w.r.t. the affine basis (a0,...,am) in E are (λ0,...,λm) with λ0 +···+λm = 1, then the coordinates of ha,λi w.r.t. the basis (a0,...,am) in Eb are  
对于任何元素ha，λi，其中，λ=06，如果w.r.t.的重心坐标，e中的仿射基（a0，…，am）是（λ0，…，λm）且λ0+···+λm=1，那么ha，λi w.r.t.eb中的基（a0，…，am）是

(λλ0,...,λλm).  
（λ0，…，λm）。

If a vector v ∈ →−E is expressed as  
如果向量v∈→−e表示为

,  
，

with respect to the affine basis (a0,...,am) in E, then its coordinates w.r.t. the basis  
关于e中的仿射基（a0，…，am），则其坐标为w.r.t.基

(a0,...,am) in Eb are  
（a0，…，am）在eb中是

(−(v1 + ··· + vm),v1,...,vm).  
（−（v1+····+vm），v1，…，vm）。

Proof. We sketch parts of the proof, leaving the details as an exercise. Figure 24.2 shows the basis () corresponding to the affine frame (a0,a1,a2) in E.  
证据。我们画出部分证据，把细节留作练习。图24.2显示了e中仿射帧（a0，a1，a2）对应的基（）。

*a*

0

*a*

1

*a*

2

x =

*a*

0

+

x

1

*a*

0

*a*

1

x

+

2

*a*

0

*a*

2

E

Ω

*a*

0

x =

E

*a*

0

*a*

2

*a*

0

*a*

1

1

x,

*λ*

x,

*λ*

*λ*

*λ*

x

1

x

2

(

,

,

)

=

Figure 24.2: The affine frame (a0,a1,a2) of E and the basis (.  
图24.2:e的仿射框（a0，a1，a2）和基（.

If we assume that we have a nontrivial linear combination  
假设我们有一个非平凡的线性组合

,  
，

if µ = 06 , then we have  
如果μ=06，那么我们有

,  
，

which is never null, and thus, µ = 0, but since () is a basis of →−E, we must also have λi = 0 for all i,1 ≤ i ≤ m.  
它从不为空，因此，μ=0，但由于（）是→−e的基础，因此我们也必须对所有i都有λi=0，1≤i≤m。

Given any element hx,λi ∈ Eb, if  
给定任何元素hx，λi∈eb，如果



over the affine frame (, in view of the definition of +b , we have  
在仿射框架上（，考虑到+b的定义，我们有

hx,λi = ha0 + x1a−−0→a1 + ··· + xma−−0a→m,λi  
hx，λi=h0+x1a−0→a1+·····+xma−0a→m，λi

= ha0,λi +b λx1a−−0→a1 +b ··· +b λxm−−a0a→m,  
=Ha0，λi+bλx1a−0→a1+b····+bλxm−a0a→m，

which shows that over the basis (, the coordinates of hx,λi are  
这表明在基础上（，hx的坐标，λi是

(λx1,...,λxm,λ).  
（λx1，…，λxm，λ）。

If (x1,...,xm) are the coordinates of x w.r.t. the affine frame ( )) in  
如果（x1，…，xm）是x w.r.t.中仿射帧（）的坐标

E, then (x1,...,xm,1) are the coordinates of x in Eb, i.e., the last coordinate is 1, and if u has coordinates (u1,...,um) with respect to the basis (, then u has coordinates (u1,...,um,0) in Eb, i.e., the last coordinate is 0. Figure 24.3 shows the affine frame (a0,a1,a2) in E viewed as a basis in Eb.  
e，那么（x1，…，xm，1）是eb中x的坐标，也就是说，最后一个坐标是1，如果u有关于基的坐标（u1，…，um），那么u在eb中有坐标（u1，…，um，0），也就是说，最后一个坐标是0。图24.3显示了作为eb基础的e中的仿射帧（a0，a1，a2）。

*a*

0

x =

E

1

x,

*λ*

x,

*λ*

*λ*

*λ*

(

,

,

)

=

*a*

0

*a*

1

*a*

2

x =

E

Ω

*λ*

*λ*

*λ*

*a*

*a*

*a*

0

0

1

1

2

2

+

+

*a*

1

Ω

*a*

2

*λ*

0

*λ*

1

*λ*

2

Figure 24.3: The basis (a0,a1,a2) in Eb.  
图24.3:eb中的基础（a0、a1、a2）。

### 24.3. ANOTHER CONSTRUCTION OF Eˆ 24.3。E\_的另一个结构

Now that we have defined Eb and investigated the relationship between affine frames in E and bases in Eb, we can give another construction of a vector space F from E and →−E that will allow us to “visualize” in a much more intuitive fashion the structure of Eb and of its operations +b and ·.  
既然我们已经定义了e b并研究了e中仿射帧与eb中碱基之间的关系，我们可以给出从e到→−e的向量空间f的另一种构造，它将使我们以更直观的方式“可视化”eb及其操作的结构+b和···。

## 24.3 Another Construction of Eb 24.3电子束的另一种结构

One would probably wish that we could start with this construction of F first, and then define Eb using the isomorphism defined below. Unfortunately, we first need the vector space structure on E to show that Ω is linear!  
有人可能希望我们可以先从F的构造开始，然后使用下面定义的同构定义eb。不幸的是，我们首先需要E上的向量空间结构来显示Ω是线性的！

Definition 24.1. Given any affine space E,, we define the vector space F as the direct sum →−E ⊕R, where R denotes the fieldF →−R⊕considered as a vector space (over itself). Denoting∈ F the unit vector in by 1, since = E R, every vector v can be written as v = u+λ1, for some unique u ∈ E and some unique λ ∈ R. Then, for any choice of an origin Ω1 in E, we define the map Ω:b Eb → F, as follows:  
定义24.1.对于任意仿射空间e，我们将向量空间f定义为直接和→−e r，其中r表示字段f→−r被视为向量空间（在其自身上）。用1表示单位向量∈f，由于=e r，每一个向量v都可以写成v=u+λ1，对于某些唯一的u∈e和一些唯一的λ∈r，那么，对于e中的原点Ω1的任何选择，我们定义了图Ω：b eb→f，如下：

∈ E and λ = 06 ;  
∈e和λ=06；

.  
.

The idea is that, once again, viewing as an affine space under its canonical structure, E is embedded in F as the hyperplane H = 1 + E, with direction , the hyperplane →−E in  
我们的想法是，再一次把e看作是一个仿射空间，在它的正则结构下，e嵌入f中，作为超平面h=1+e，有方向，超平面→−e嵌入

F. Then, every point a ∈ E is in bijection with the point A = 1 + Ω1a, in the hyperplane H. If we denote the origin 0 of the canonical affine space F by Ω, the map Ω maps a pointb ha,λi ∈ E to a point inF F, as follows: ) is the point on the line passing through both→− the origin Ω of and the point A = 1 + Ω1a in the hyperplane H = 1 + E, such that  
f.那么，在超平面h中，每个点a∈e都是双射的，点a=1+Ω1a。如果我们用Ω来表示正则仿射空间f的原点0，那么mapΩ将A点b ha，λi∈e映射到一个点inf f，如下所示：）是通过这两个点的直线上的点→−原点Ω。超平面h=1+e中的点a=1+Ω1a，这样

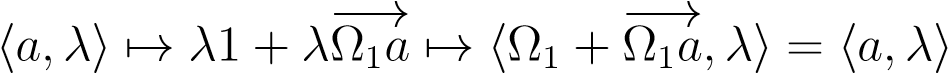
.  
.

The following proposition shows that Ω is an isomorphism of vector spaces.  
下面的命题表明，Ω是向量空间的同构。

Proposition 24.4. Given any affine space (E,→−E), for any choice Ω1 of an origin in E, the map Ω:b Eb → F is a linear isomorphism between Eb and the vector space F of Definition 24.1. The inverse of Ωb is given by  
提案24.4.给定任意仿射空间（e，→−e），对于e中原点的Ω1的任何选择，mapΩ：b eb→f是eb与定义为24.1的向量空间f之间的线性同构。Ωb的倒数由下式得出：

; .  
；

Proof. It is a straightforward verification. We check that Ω is invertible, leaving the verifi-b cation that it is linear as an exercise. We have  
证据。这是一个简单的验证。我们检查了Ω是可逆的，剩下的校验B阳离子是线性的作为练习。我们有



and  
和

,  
，

and since Ω is the identity onb →−E, we have shown that Ω ◦ Ω−1 = id, and Ω = id. This shows that Ω is a bijection.b   
既然Ω是B→−E上的标识，我们已经证明了ΩΩ−1=ID，Ω=ID。这表明Ω是双射。b

Figure 24.4 illustrates the embedding of the affine space E into the vector space F, when E is an affine plane.  
图24.4说明了当e是仿射平面时，仿射空间e嵌入向量空间f。

Ω

*E*

Ω

1

*a*

*A*

=

1

+

Ω

1

*a*

*λ*

Ω

*A*

*H*

=

1

+

*E*

1

Figure 24.4: Embedding an affine space E, into a vector space F.  
图24.4：在向量空间f中嵌入仿射空间e。

Proposition 24.4 gives a nice interpretation of the sum operation +b of Eb. Given two weighted points ha1,λ1i and ha2,λ2i, we have  
提案24.4对电子商务的和运算+b给出了一个很好的解释。给定两个加权点ha1，λ1i和ha2，λ2i，我们得到

ha1,λ1i +b ha2,λ2i = Ωb−1(Ω(b ha1,λ1i) + Ω(b ha2,λ2i)).  
Ha1，λ1i+b Ha2，λ2i=Ωb−1（Ω（b Ha1，λ1i）+Ω（b Ha2，λ2i））。

The operation) has a simple geometric interpretation. If λ1 +λ2 = 06 , then find the points M1 and M2 on the lines passing through the origin Ω of−−→ F→and the points−−→  
操作）具有简单的几何解释。如果λ1+λ2=06，则在穿过−→F→原点Ω的直线上找到点M1和M2，然后找到点−→

) and A2 = Ω(b a2) in the hyperplane H, such that ΩM1 = λ1Ω−−A1 and ΩM2 = λ2ΩA2, add the vectors Ω−−M→1 and , getting a point N such that Ω−−→N = −−ΩM→1 + Ω−−M→2, and consider the intersection G of the line passing through Ω and N with the hyperplane H. Then, G is the barycenter of A1 and A2 assigned the weights λ1/(λ1 +λ2) and λ2/(λ1 +λ2), and if g = Ωb−1(Ω−→G), then Ωb−1(Ω−−→N) = hg,λ1 + λ2i. See Figure 24.5.  
）在超平面H中，A2=Ω（b a2），这样，Ωm 1=λ1Ω−a1和Ωm 2=λ2Ωa2，加上矢量Ω−m→1，得到一个点n，使得Ω−→n=−Ωm→1+Ω−m→2，并考虑通过Ω和n的线与超平面的交点g。e h.那么，g是分配重量为λ1/（λ1+λ2）和λ2/（λ1+λ2）的a1和a2的重心，如果g=Ωb−1（Ω−→g），则Ωb−1（Ω−→n）=hg，λ1+λ2i。见图24.5。

Instead of adding the vectors and , we can take the middle N0 of the segment M1M2, and G is the intersection of the line passing through Ω and N0 with the hyperplane H as illustrated in Figure 24.5.  
不用加向量和，我们可以取段m1m2的中间n0，g是穿过Ω和n0的线与超平面h的交点，如图24.5所示。

### 24.3. ANOTHER CONSTRUCTION OF Eˆ 24.3。E\_的另一个结构

*E*

Ω

1

*a*

*A*

=

1

+

Ω

1

*a*

*λ*

Ω

*A*

*H*

=

1

+

*E*

1

1

1

1

1

1

Ω

*a*

2

2

1

Ω

*A*

=

1

+

Ω

1

*a*

2

2

*λ*

Ω

*A*

2

*E*

*A*

Ω

1

Ω

*A*

2

M

1

M

*Ω*

M

1

*Ω*

M

2

N

N

’

2

N

2

N

*E*

*A*

*H*

=

1

+

*E*

1

Ω

*A*

2

M

N

’

M

1

G

Figure 24.5: The geometric construction of Ω(b ha1,λ1i) + Ω(b ha2,λ2i) for λ1 + λ2 = 0.6  
图24.5：λ1+λ2=0.6的Ω（b ha1，λ1i）+Ω（b ha2，λ2i）的几何结构

If λ1 + λ2 = 0, then ha1,λ1i +b ha2,λ2i is a vector determined as follows. Again, find the points M1 and M2 on the lines passing through the origin Ω of F and the points  
如果λ1+λ2=0，则Ha1、λ1i+b Ha2、λ2i是如下确定的向量。同样，在穿过F原点Ω的直线上找到M1和M2点以及这些点

and A2 = Ω(b a2) in the hyperplane H, such that and, and add the vectors Ω−−M→1 and , getting a point N such that ΩN = ΩM1 + ΩM2. The desired vector is Ω−−→N, which is parallel to the line A1A2. Equivalently, let N0 be the middle of the segment M1M2, and the desired vector is 2Ω−−N→0. See Figure 24.6.  
在超平面h中，a2=Ω（ba2），这样，and，并加上矢量Ω−m→1，得到点n，这样，Ωn=Ωm 1+Ωm2。所需的矢量为Ω−→N，与管线A1A2平行。等价地，让n0是段m1m2的中间，并且所需的向量是2Ω−n→0。见图24.6。

We can also give a geometric interpretation of ha,λi+u. Let) in the hyperplane H, let D be the line determined by A and u, let M1 be the point such that , and let M2 be the point such that Ω−−M→2 = u, that is, M2 = Ω+u. By construction, the line D is in the hyperplane H, and it is parallel to Ω−−M→2, so that D, M1, and M2 are coplanar. Then, add the vectors Ω−−M→1 and , getting a point N such that Ω−−→N = Ω−−M→1 + Ω−−M→2, and let G be the intersection of the line determined by Ω and N with the line D. If g = Ωb−1(Ω−→G), then,. Equivalently, if N0 is the middle of the segment M1M2, then G is the intersection of the line determined by Ω and N0, with the line D; see Figure 24.7.  
我们还可以给出超平面h中h a，λi+u.let）的几何解释，设d为a和u确定的线，设m1为点，使m2为点，使Ω−−m→2=u，即m2=Ω+u。通过构造，d线在超平面h中，它是Parall。EL至Ω−m→2，使d、m1和m2共面。然后，加上矢量Ω−m→1，得到一个点n，使Ω−→n=Ω−m→1+Ω−m→2，并让g是由Ω和n确定的线与线d的交点。如果g=Ωb−1（Ω−→g），则，。同样，如果n0是段m1m2的中间，则g是由Ω和n0确定的线与线d的交点；见图24.7。

We now consider the universal property of Eb mentioned at the beginning of this section.  
我们现在考虑本节开头提到的电子商务的普遍属性。

*E*

Ω

1

*a*

*A*

=

1

+

Ω

1

*a*

*λ*

Ω

*A*

*H*

=

1

+

*E*

1

1

1

1

1

1

Ω

*a*

2

2

1

Ω

*A*

=

1

+

Ω

1

*a*

2

2

*λ*

Ω

2

*E*

*A*

Ω

1

Ω

M

1

M

*Ω*

M

1

*Ω*

M

2

N

N

’

2

*A*

N

*A*

2

Figure 24.6: The geometric construction of Ω(b ha1,λ1i) + Ω(b ha2,λ2i) for λ1 + λ2 = 0.  
图24.6：λ1+λ2=0时，Ω（b ha1，λ1i）+Ω（b ha2，λ2i）的几何结构。

## 24.4 Extending Affine Maps to Linear Maps 24.4将仿射映射扩展到线性映射

Roughly, the vector space Eb has the property that for any vector space →−F and any affine map f : E → →−F , there is a unique linear map extending f : E → →−F . As a consequence, given two affine spaces E and F, every affine map f : E → F extends uniquely to a linear map fb: Eb → Fb. First, we define rigorously the notion of homogenization of an affine space.  
大体上，向量空间eb具有以下性质：对于任何向量空间→−f和任何仿射映射f:e→→−f，都有一个唯一的线性映射扩展f:e→→−f。因此，在给定两个仿射空间e和f的情况下，每个仿射映射f:e→f唯一地扩展到一个线性映射fb:eb→fb。首先，我们严格定义了仿射空间的同质化概念。

Definition 24.2. Given any affine space E,, a homogenization (or linearization) of  
定义24.2.给定任意仿射空间e，，的均匀化（或线性化）

(E,→−E) is a triple hE,j,ωi, where E is a vector space,→− → E j:E →E → E→− is an injective affine map with associated injective linear map i: E , ω: R is a linear form such that ), and for every vector space F and every affine map f : E → F there is a unique linear map extending f, i.e., f = fb◦ j, as in the following diagram:  
（e，→−e）是一个三重He，j，ωi，其中e是一个向量空间，→–→e j:e→e→e→−是一个与之相关的内射线性映射的内射仿射映射i:e，ω：r是一个这样的线性形式），对于每个向量空间f和每个仿射映射f:e→f，都有一个唯一的线性映射扩展如下图所示：

j  
J

E@@@@@@@@ / E fb  
E@@@@@@@@@@/E FB

f  
f

→−F  
→−F

Thus, j(E) = ω−1(1) is an affine hyperplane with direction (0). Note that we could have defined a homogenization of an affine space (E, E), as a triple hE,j,Hi, where E is a vector space, H is an affine hyperplane in E, and j: E → E is an injective affine map such that j(E) = H, and such that the universal property stated above holds. However, Definition 24.2 is more convenient for our purposes, since it makes the notion of weight more evident.  
因此，j（e）=ω−1（1）是一个方向为（0）的仿射超平面。注意，我们可以定义一个仿射空间（e，e）的同构，作为一个三重He，j，hi，其中e是一个向量空间，h是e中的仿射超平面，j:e→e是一个可注射的仿射映射，这样j（e）=h，并且上面所述的普遍性质成立。然而，定义24.2对于我们的目的来说更方便，因为它使重量的概念更加明显。

The obvious candidate for E is the vector space Eb that we just constructed. The next proposition will show that Eb indeed has the required extension property. As usual, objects  
显然，e的候选者是我们刚刚构造的向量空间eb。下一个提议将表明电子商务确实具有所需的扩展属性。像往常一样，物体

### 24.4. EXTENDING AFFINE MAPS TO LINEAR MAPS

*E*

Ω

1

*a*

*A*

=

1

+

Ω

1

*a*

*λ*

Ω

*A*

*H*

=

1

+

*E*

1

Ω

u

=

M

1

D

M

2

*E*

*A*

*H*

=

1

+

*E*

1

Ω

u

1

D

M

2

M

*E*

*A*

Ω

u

1

D

M

2

M

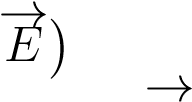
N

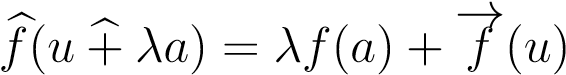
G

N’

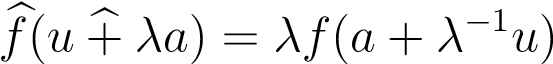
Figure 24.7: The geometric construction of h*a,λ*i + *u*.

defined by a universal property are unique up to isomorphism. This property is left as an exercise.

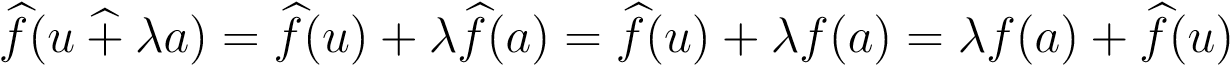
**Proposition 24.5.** *Given any affine space E,* *and any vector space* →−*F , for any affine map f* : *E* → →−*F , there is a unique linear map*  *extending f such that*



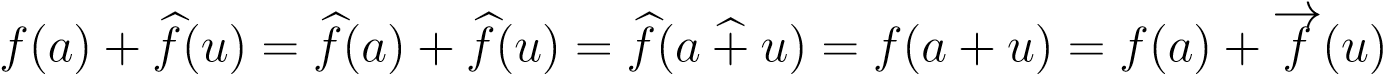
*for all a* ∈ *E, all u* ∈ →−*E, and all λ* ∈ R*, where* →−*f is the linear map associated with f. In particular, when λ* = 06 *, we have*

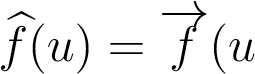
*.*

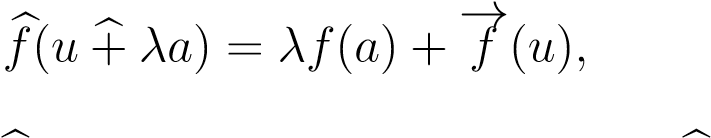
*Proof.* Assuming that *f*b exists, recall that from Proposition 24.1, for every *a* ∈ *E*, every element of *E*b can be written uniquely as *u*+b *λa*. By linearity of *f*band since *f*bextends *f*, we have

*.*

If *λ* = 1, since  and *a* + *u* are identified, and since *f*bextends *f*, we must have

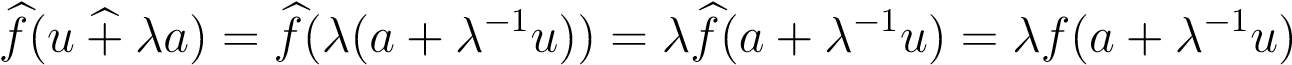
*,*

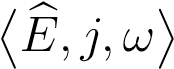
and thus) for all *u* ∈ →−*E*. Then we have



which proves the uniqueness of *f*. On the other hand, the map *f* defined as above is clearly a linear map extending *f*.

When *λ* = 06 , we have

*.*

Proposition 24.5 shows that , is a homogenization of *E,*. As a corollary, we obtain the following proposition.

**Proposition 24.6.** *Given two affine spaces E and F and an affine map f* : *E* → *F, there is a unique linear map**extending f, as in the diagram below,*

*f*

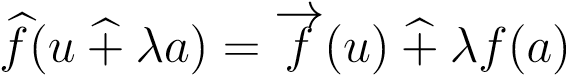
*E* / *F*

*jj*

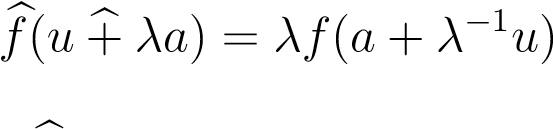
*E*b / *F*b

*f*b

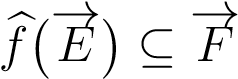
*such that*

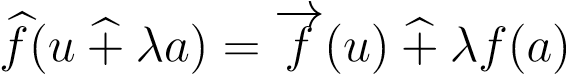
*,*

*for all a* ∈ *E, all u* ∈ →−*E, and all λ* ∈ R*, where* →−*f is the linear map associated with f. In particular, when λ* = 06 *, we have*

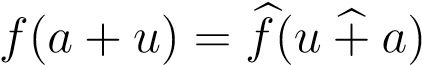
*.*

*Proof.* Consider the vector space and the affine map *j* ◦ *f* : *E* → *F*b. By Proposition 24.5, there is a unique linear map *f* : *E* → *F* extending *j* ◦ *f*, and thus extending *f*.

Note that *f*b: *E*b → *F*b has the property that. More generally, since

*,*

the linear map *f*b is weight-preserving. Also observe that we recover *f* from *f*b, by letting *λ* = 1 in *f*b(*u* +b *λa*) = *λf*(*a* + *λ*−1*u*), that is, we have

*.*